

---

# **BACHELORARBEIT**

---

Herr  
**Jens Schreiter**

**Methoden zur Sensitivitäts-  
und Robustheitsbewertung  
für multikriterielle  
Optimierungsverfahren**

2010



# **Bachelorarbeit**

---

## **Methoden zur Sensitivitäts- und Robustheitsbewertung für multikriterielle Optimierungsverfahren**

Autor:

**Jens Schreiter**

Studiengang:

Angewandte Mathematik

Seminargruppe:

MA07w1-B

Erstprüfer:

Prof. Dr. rer. nat. Ullrich Griesbach

Zweitprüfer:

Dipl. Math. (FH) Steffen Kux

Mittweida, August 2010



## **Bibliographische Angaben**

Schreiter, Jens: Methoden zur Sensitivitäts- und Robustheitsbewertung für multikriterielle Optimierungsverfahren, 61 Seiten mit Anhang, 29 Abbildungen, 7 Tabellen, Hochschule Mittweida (FH), Fakultät Mathematik/Naturwissenschaften/Informatik

Bachelorarbeit, 2010

## **Referat**

Die Bachelorarbeit beschäftigt sich einerseits mit verschiedenen Verfahren zur Sensitivitätsanalyse für multikriterielle Optimierungsverfahren. Die Sensitivität wird dabei a-priori und a-posteriori untersucht. Andererseits werden Methoden zur Bewertung der Robustheit von Lösungen aus multikriteriellen Optimierungsverfahren vorgestellt. Zudem werden die Ergebnisse der Sensitivitätsanalyse und Robustheitsbewertung in geeigneten graphischen Darstellungen visualisiert. Abschließend werden die erläuterten Analyse- und Bewertungsmethoden auf einen Praxisbeispiel angewendet und miteinander verglichen.



# Inhaltsverzeichnis

<b>Abbildungsverzeichnis</b>	<b>v</b>
<b>Tabellenverzeichnis</b>	<b>vii</b>
<b>Abkürzungsverzeichnis</b>	<b>ix</b>
<b>Notationen</b>	<b>xi</b>
<b>Vorwort</b>	<b>xiii</b>
<b>1 Einleitung</b>	<b>1</b>
1.1 Motivation . . . . .	1
1.2 Die IAV-EngineeringToolbox . . . . .	1
1.3 Ziel der Arbeit . . . . .	3
<b>2 Mathematische Grundlagen</b>	<b>5</b>
2.1 Multikriterielle Optimierung . . . . .	5
2.2 Stochastik . . . . .	7
2.2.1 Wahrscheinlichkeitstheorie . . . . .	7
2.2.2 Deskriptive Statistik . . . . .	11
2.2.3 Multiple lineare Regression . . . . .	12
<b>3 Sensitivitätsanalyse</b>	<b>17</b>
3.1 A-priori Sensitivität . . . . .	17
3.2 A-posteriori Sensitivität . . . . .	21
3.2.1 Regression und Korrelation . . . . .	21
3.2.2 Methode nach Kendall . . . . .	26
3.2.3 Bewertung mit ANOVA . . . . .	27
<b>4 Robustheitsbewertung</b>	<b>29</b>
4.1 Direkte Auswertung . . . . .	29
4.1.1 Bewertung mittels Streuungsmaße . . . . .	30
4.1.2 Bewertung durch Prognose der Zielfunktionswertänderungen . . . . .	32
4.2 Aufbauende Verfahren . . . . .	34
4.3 Robuste Optimierung . . . . .	35
4.4 Bewertungsablauf . . . . .	36
<b>5 Technisches Anwendungsbeispiel</b>	<b>39</b>
5.1 Problemstellung . . . . .	39

---

5.2	Auswertung . . . . .	42
5.2.1	Sensitivitätsanalyse . . . . .	42
5.2.2	Robuste Lösungen . . . . .	45
<b>6</b>	<b>Schluss</b>	<b>51</b>
6.1	Zusammenfassung . . . . .	51
6.2	Ausblick . . . . .	52
<b>A</b>	<b>Anhang</b>	<b>53</b>
A.1	Verwendete Formeln . . . . .	53
A.1.1	Wichtige Verteilungen . . . . .	53
A.1.2	Statistische Lageparameter und Streuungsmaße . . . . .	53
A.2	Graphische Darstellungen zum Anwendungsbeispiel . . . . .	55
	<b>Literaturverzeichnis</b>	<b>59</b>
	<b>Erklärung</b>	<b>61</b>



# Abbildungsverzeichnis

1.1 Die EngineeringToolbox mit dem Modul MCA . . . . .	2
2.1 Paretofront einer MOA mit zwei Zielfunktionen . . . . .	7
2.2 Dichtefunktionen zur Schiefe . . . . .	9
2.3 Dichtefunktionen zum Exzess . . . . .	9
3.1 Modifizierte Branin-Funktion im Bereich $[6, 13] \times [-5, 5]$ . . . . .	20
3.2 Modifizierte Branin-Funktion im Bereich $[-6, 21] \times [-12, 15]$ . . . . .	20
3.3 Balkendiagramme der Beiträge zum Zielfunktionswert in $\mathbf{x}^{(1)}$ und $\mathbf{x}^{(2)}$ . . . . .	21
3.4 Linearer und quadratischer Zusammenhang . . . . .	22
3.5 Balkendiagramm mit Korrelationskoeffizienten . . . . .	23
3.6 Korrelationsdiagramm . . . . .	24
3.7 Auswirkung von Ausreißern auf die Korrelationskoeffizienten $r_{YX}$ und $k_{YX}$ . . . . .	25
4.1 Zu minimierende Zielfunktion mit globalen Optimum $x^*$ und robusten Optimum $x_R$ . . . . .	29
4.2 Histogramme zur empirischen Verteilung der Zielgrößen . . . . .	30
4.3 Beispielhafte Darstellung der Variationsverhältnisse $\psi_{Y_i X_j}$ einer Lösung $\mathbf{x}$ einer MOA . . . . .	32
4.4 Graphische Darstellung der Zielfunktionswertänderungen bei Variation einer Lösung $\mathbf{x}$ . . . . .	33
4.5 Balkendiagramm der Robustheitsgrade $\eta_i$ einer Lösung $\mathbf{x}$ aus einer MOA . . . . .	34
4.6 Monte-Carlo Simulation zweier normalverteilter Zufallsgrößen . . . . .	35
5.1 Simulation von Steuerketten durch V-CD . . . . .	39
5.2 Geometrische Darstellung der Parameter des Steuertriebes . . . . .	40
5.3 Korrelationsdiagramm zum Anwendungsbeispiel . . . . .	43
5.4 Korrelationskoeffizienten zwischen der Zielgröße H und allen Parametern . . . . .	43
5.5 Entscheidungsgraph mit den Lösungen $\mathbf{x}_1$ , $\mathbf{x}_2$ und $\mathbf{x}_3$ . . . . .	45
5.6 Variationsverhältnisse $\psi$ der Lösungen $\mathbf{x}_1$ und $\mathbf{x}_3$ . . . . .	46
5.7 Zielfunktionswertveränderungen der drei Lösungen . . . . .	48
A.1 Korrelationskoeffizienten zwischen der Zielgröße V und allen Parametern . . . . .	55
A.2 Korrelationskoeffizienten zwischen der Zielgröße $V_v$ und allen Parametern . . . . .	56
A.3 Korrelationskoeffizienten zwischen der Zielgröße T und allen Parametern . . . . .	56
A.4 Korrelationskoeffizienten zwischen der Zielgröße $P_r$ und allen Parametern . . . . .	57
A.5 Korrelationskoeffizienten zwischen der Zielgröße U und allen Parametern . . . . .	58



# Tabellenverzeichnis

5.1 Beschreibung der Parameter des Anwendungsbeispieles . . . . .	41
5.2 Beschreibung der Ziele des Anwendungsbeispieles . . . . .	41
5.3 Bestimmtheitsmaße der multiplen Regressionen . . . . .	42
5.4 Zusammenfassung der Sensitivitätsanalyse des Anwendungsbeispieles . . . . .	44
5.5 Zielfunktionswerte der ausgewählten Lösungen . . . . .	45
5.6 Robustheitsbewertung mittels Streuungsmaße . . . . .	46
5.7 Robustheitsbewertung mittels Zielfunktionswertänderungen . . . . .	47



# Abkürzungsverzeichnis

Abb.	Abbildung
ANOVA	Analysis Of Variants
CAD	Computer Aided Design
CAE	Computer Aided Engineering
ET	Engineering Toolbox
IAV	Ingenieurgesellschaft für Auto und Verkehr
IGIS	lineares Gleichungssystem
MCA	Multi Criteria Analysis
MOA	multikriterielle Optimierungsaufgabe
MrBOA	Multiobjective real-coded Bayesian Optimization Algorithm
o.B.d.A.	ohne Beschränkung der Allgemeinheit
Tab.	Tabelle
UNO	Universal Optimization
V-CD	Virtual Chain Drive
ZG	Zufallsgröße



# Notationen

$\emptyset$	Leere Menge
$\mathbb{N}$	Menge der natürlichen Zahlen $\{1, 2, 3, \dots\}$
$\mathbb{Z}$	Menge der ganzen Zahlen $\{\dots, -2, -1, 0, 1, 2, \dots\}$
$\mathbb{R}$	Menge der reellen Zahlen
$ x $	Betrag von $x \in \mathbb{R}$
$\log(x)$	Logarithmus von $x$ zur Basis 10
$\delta_{ij}$	Kronecker Symbol $\begin{cases} 1, & \text{falls } i = j \\ 0, & \text{sonst.} \end{cases}$
$\mathbf{x}$	Vektor
$\nabla f(\mathbf{x})$	Gradient von $f(\mathbf{x})$
$\mathbf{f}(\mathbf{x})$	Vektorfunktion
$\mathbf{A}$	Matrix
$\mathbf{A}^T$	Transponierte Matrix zu $\mathbf{A}$
$\det(\mathbf{A})$	Determinante von $\mathbf{A}$
$\mathbf{A}^{-1}$	Inverse Matrix zu $\mathbf{A}$
$\mathcal{X}$	Zufallsgröße
$\varphi(x)$	Dichtefunktion der standardisierten Normalverteilung
$\Phi(x)$	Verteilungsfunktion der standardisierten Normalverteilung
$\psi$	Variationsverhältnis
$\iota$	Robustheitsindex
$\eta$	Robustheitsgrad





# Vorwort

Diese Bachelorarbeit zum Thema „Methoden zur Sensitivitäts- und Robustheitsbewertung für multikriterielle Optimierungsverfahren“ entstand im Rahmen des Studiums an der Hochschule Mittweida (FH) im Studiengang Angewandte Mathematik und in Zusammenarbeit mit der IAV GmbH am Entwicklungsstandort in Chemnitz.

Für die angenehme Zusammenarbeit während meines gesamten Bachelorprojektes danke ich allen Mitarbeitern der Abteilung DP-P3 der IAV GmbH. Mein besonderer Dank gilt dabei meinen Betreuer Dipl. Math. (FH) Steffen Kux für die zahlreichen Ideen zur Umsetzung und Ausarbeitung des Themas. Für die Unterstützung bei der programmiertechnischen Umsetzung sei Herrn Dipl. Wirt.-Inf. Michael Meyer gedankt. Einen Dank gilt auch Dipl. Math. Martin Stöcker für die inspirierenden Gespräche und der Begutachtung der Arbeit. Gedankt sei weiterhin dem Abteilungsleiter Dr. Horst Liebermann für die konstruktiven Ratschläge zur Erweiterung der in dieser Arbeit vorgestellten Methoden auf ähnliche Sachverhalte. Zu guter Letzt möchte ich meinen betreuenden Professor Dr. rer. nat. Ullrich Griesbach danken, der wesentliche Verbesserungsvorschläge eingebracht und den erfolgreichen Abschluss der Arbeit ermöglicht hat.

Chemnitz, 30. August 2010

Jens Schreiter



# 1 Einleitung

## 1.1 Motivation

In der heutigen Wissenschaft und Technik sind Problemstellungen aus der Optimierung weit verbreitet. Das Spektrum reicht hierbei von Anwendungen in der Volkswirtschaft bis hin zu den unterschiedlichsten Ingenieurwissenschaften. Durch den technischen Fortschritt und den immer kürzer werdenden Entwicklungszeiten wird der Einsatz von Optimierungsverfahren immer wichtiger. Zudem werden die Optimierungsaufgaben immer umfangreicher und komplexer. Dadurch sind viele moderne Optimierungsverfahren entstanden, die meistens auf biologischen oder physikalischen Sachverhalten beruhen. Die umfangreichen und vielseitigen Problemstellungen erschweren zusätzlich das Ausmachen der wichtigsten Parameter. Auch die Einflüsse von Parametern auf die Zielgrößen sind dadurch schwer zu erkennen. Bei der Beurteilung der Optimierungsergebnisse ist von großer Bedeutung, wie robust die Ergebnisse sind, da die erforderlichen Eigenschaften der Lösungen auch dann gewährleistet sein müssen, wenn die Parameter kleinen Schwankungen, wie z.B. Toleranzen, unterworfen sind. Schließlich hat die Robustheit der berechneten Ergebnisse entscheidenden Einfluss auf die Qualität und Funktionsweise der optimierten Produkte.

In der Ingenieurgesellschaft für Auto und Verkehr, kurz IAV GmbH, einem weltweit vertretenen und führenden Engineering-Partner der Automobilindustrie, entstehen solche Fragestellungen aus praxisrelevanten Optimierungsaufgaben. Beispielsweise bei der Auslegung von Antriebssträngen, Motorenentwicklung oder allgemein bei der internen Fahrzeugentwicklung entstehen komplexe Optimierungsprobleme, die gelöst und ausgewertet werden müssen. Deshalb existiert in der IAV die EngineeringToolbox, in dem neben verschiedenen CAE-Anwendungen ein Modul zur Lösung von verschiedenen Optimierungsaufgaben bereit steht. Zur Auswertung und Visualisierung der aus der Optimierung erhaltenen Ergebnisse hält die IAV EngineeringToolbox ein weiteres Modul bereit.

## 1.2 Die IAV-EngineeringToolbox

Die IAV-EngineeringToolbox ist ein umfangreicher Werkzeugkasten von CAE-Anwendungen, der den Ingenieuren der IAV zur Verfügung gestellt wird. Neben verschiedenen Tools zur Fahrzeugentwicklung beinhaltet die ET die Module UNO und MCA zur Optimierung und Entscheidungsunterstützung. Im UNO sind unterschiedliche Verfahren, z.B. Gradientenverfahren, stochastische

Suchen, Genetische Algorithmen, Ersatzmodell-gestützte Verfahren, Partikelschwarm Algorithmen und weitere zur Lösung diskreter, nichtlinearer, multikriterieller und anderer Optimierungsprobleme enthalten. Für bessere Performance sind die oben genannten Verfahren hinsichtlich der Berechnung von Zielfunktionswerten parallelisiert worden. Besonders bei Optimierungsproblemen mit rechenzeitintensiven Zielfunktionsauswertungen bringt die Parallelisierung eine enorme Zeitersparnis. In den Entwicklungsbereichen der IAV GmbH entstehen oft multikriterielle Optimierungsaufgaben. Um die daraus erhaltenen Lösungsmengen zu analysieren und praxisrelevante Lösungen herausfiltern, ist eigens in der IAV GmbH das Modul MCA entwickelt worden. In Abb. 1.1 ist das Modul MCA mit einem Standardbeispiel dargestellt. Im Modul MCA sind neben

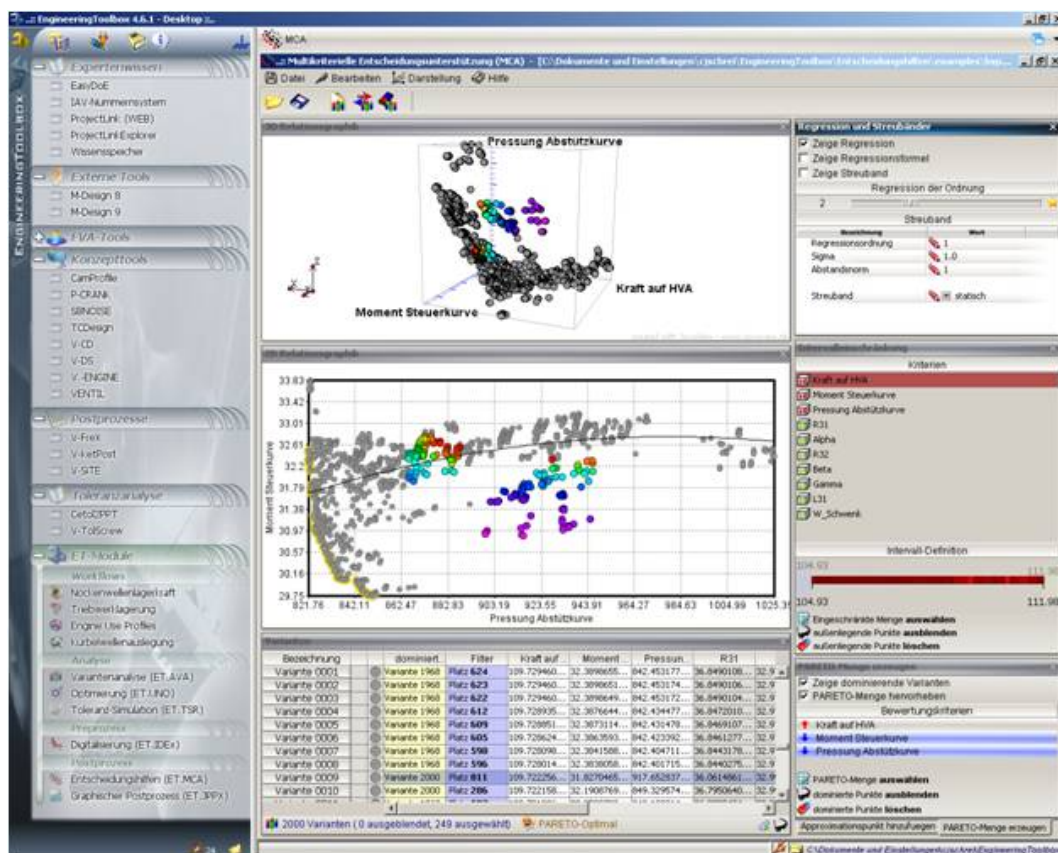


Abb. 1.1: Die EngineeringToolbox mit dem Modul MCA

verschiedenen graphischen Darstellungen zur Auswertung der erhaltenen Lösungsmenge auch Möglichkeiten zur statistischen Analyse der Lösungsdaten integriert. Die wichtigsten Elemente zur Veranschaulichung der Ergebnisse im Modul MCA sind die 2D- und 3D-Relationsgraphiken, worin eine Punktwolke der jeweiligen Dimension zu sehen ist (siehe Abb. 1.1).

### 1.3 Ziel der Arbeit

Ziel dieser Arbeit ist die Untersuchung von Verfahren zur Auswertung der aus einer Optimierung stammenden Lösungsmenge hinsichtlich Sensitivität und Robustheit. Näheres zur Sensitivitätsanalyse wird im Kapitel 3 erläutert. Im Kapitel 4 werden verschiedene Methoden zur Robustheitsbewertung betrachtet und verglichen. Parallel zur Vorstellung der Verfahren für die Sensitivitäts- und Robustheitsbewertung sollen verschiedene Darstellungen zur Visualisierung der Ergebnisse betrachtet werden. Dabei gilt es, geeignete Methoden zur Berechnung und Veranschaulichung der Ergebnisse aus der Sensitivitäts- und Robustheitsbewertung im Modul MCA der IAV-EngineeringToolbox zu integrieren. Abschließend wird ein Beispiel aus der Praxis mit den vorgestellten Methoden bezüglich Sensitivität und Robustheit ausgewertet und die Ergebnisse gegenübergestellt.



## 2 Mathematische Grundlagen

In diesem Kapitel werden benötigte Grundlagen aus der Mathematik erläutert. Zu Beginn werden Definitionen aus der multikriteriellen Optimierung vorgestellt. Anschließend werden Kenntnisse aus der Wahrscheinlichkeitsrechnung und mathematischen Statistik beschrieben.

### 2.1 Multikriterielle Optimierung

Als Erstes wird die multikriterielle Optimierungsaufgabe, kurz MOA, definiert. Im Folgenden werden nur zu minimierende Zielfunktionen betrachtet, da eine Maximierung einer Funktion  $f$  äquivalent zur Minimierung von  $-f$  ist.

**Definition 2.1:** MOA

Seien  $f_i(\mathbf{x}) : \mathbb{R}^n \rightarrow \mathbb{R}$  die Zielfunktionen für  $i = 1, \dots, m$  mit  $\mathbf{x} \in \mathbb{R}^n$ . Gesucht ist

$$\min_{\mathbf{x} \in R} f_i(\mathbf{x}) \quad i = 1, \dots, m$$

mit den durch Ungleichungen und Gleichungen

$$\begin{aligned} g_j(\mathbf{x}) &\leq 0 & j = 1, \dots, p \\ h_k(\mathbf{x}) &= 0 & k = 1, \dots, q \end{aligned}$$

beschriebenen Restriktionsbereich  $R$ . Ist ein Punkt  $\mathbf{x} \in R \subseteq \mathbb{R}^n$ , dann heißt er zulässig. Außerdem wird vorausgesetzt, dass die MOA lösbar ist. Insbesondere gilt  $R \neq \emptyset$ .

Die einzelnen Zielfunktionen werden in dieser Arbeit auch Zielgrößen genannt. Die Komponenten von  $\mathbf{x} = (x_1, \dots, x_n)^T$  heißen hierbei Variablen, Parameter und später auch Einflussgrößen. Weiterhin wird davon ausgegangen, dass die MOA nichtlinear ist. Gilt  $m = 1$ , so wird die MOA auch als einkriterielle Optimierungsaufgabe bezeichnet. Der wesentliche Unterschied zwischen diesen beiden Optimierungsproblemen besteht in der Vielfalt der Lösungsmenge. Während bei einkriteriellen Aufgabenstellungen eine optimale Lösung existiert, entstehen bei MOA's Kompromisslösungen, da die unterschiedlichen Zielfunktionen im Allgemeinen untereinander konkurrieren. Beispielsweise tritt ein Konflikt zwischen zwei Zielfunktionen auf, wenn die Leistung eines Ottomotors erhöht und gleichzeitig der Verbrauch gesenkt werden soll.

Zur Beschreibung der Kompromissmenge, die im Weiteren auch als Paretomenge bezeichnet wird, werden folgende Definitionen benötigt:

**Definition 2.2:** Dominanz

Es sei eine MOA gegeben. Ein Punkt  $\mathbf{x} \in R$  dominiert  $\mathbf{y} \in R$ , im Zeichen  $\mathbf{x} \prec \mathbf{y}$ , wenn

$$f_i(\mathbf{x}) \leq f_i(\mathbf{y}) \quad \forall i \in \{1, \dots, m\} \quad \wedge \quad \exists j \in \{1, \dots, m\} \quad \text{mit} \quad f_j(\mathbf{x}) < f_j(\mathbf{y}).$$

gilt. Ein zulässiger Punkt  $\mathbf{x} \in R$  heißt genau dann nichtdominiert bezüglich einer Menge  $Y \subseteq R$ , wenn  $\nexists \mathbf{y} \in Y$ , für das  $\mathbf{y} \prec \mathbf{x}$  gilt. Dann schreiben wir  $\mathbf{x} \prec Y$ .

Wenn eine Lösung eine andere dominiert, bedeutet das, dass keiner ihrer Zielfunktionswerte schlechter und mindestens einer besser ist.

**Bemerkung 2.1:**

Hierbei können drei verschiedene Fälle auftreten, und zwar:

1.  $\mathbf{x} \prec \mathbf{y}$ ,
2.  $\mathbf{y} \prec \mathbf{x}$  oder
3. keiner von beiden dominiert den jeweils anderen.

Von besonderer Bedeutung sind diejenigen Lösungen einer MOA, die von keiner anderen dominiert werden.

**Definition 2.3:** paretooptimal, Paretomenge, Paretofront

Ein zulässiger Punkt  $\mathbf{p}$  einer MOA heißt paretooptimal, wenn  $\mathbf{p} \prec R$  gilt. Die Gesamtheit aller paretooptimalen Punkte  $\mathbf{p}$  wird zur Paretomenge  $P$  zusammengefasst. Sei  $P \subset R$  die Paretomenge einer MOA. Die Menge

$$P_f = \left\{ \mathbf{f}(\mathbf{p}) = (f_1(\mathbf{p}), \dots, f_m(\mathbf{p}))^T \mid \mathbf{p} \in P \right\}$$

heißt Paretofront.

Oft werden paretooptimale Punkte auch als effiziente Punkte oder Kompromisslösungen bezeichnet. In Abbildung 2.1 ist die Paretofront zweier zu minimierender Zielfunktionen dargestellt. In hellblau ist darin das Bild des Restriktionsbereichs  $R$  zu sehen. Das Bild von  $R$  ist hier zusammenhängend dargestellt, was nicht unbedingt der Fall sein muss. Zudem kann die Paretofront einer MOA nicht zusammenhängend sein.

Verfahren zur Lösung multikriterieller Optimierungsaufgaben werden in dieser Arbeit nicht vorgestellt, da sich die nachfolgenden Kapitel nur auf die Auswertung der erhaltenen Ergebnisse beziehen. Einen guten Überblick über die Einteilung von Optimierungsverfahren liefert [12], wobei auch einzelne Algorithmen detailliert beschrieben werden. Weitere Verfahren werden in [2] und [11] vorgestellt.



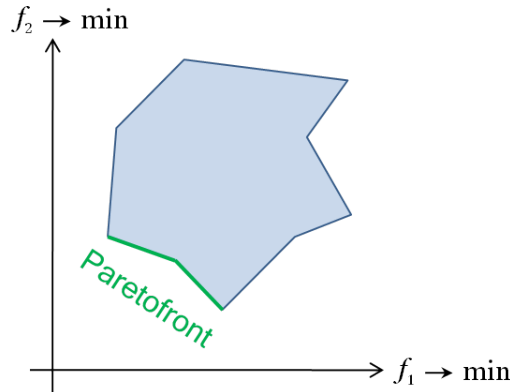


Abb. 2.1: Paretofront einer MOA mit zwei Zielfunktionen

## 2.2 Stochastik

Hier werden Grundlagen aus der Wahrscheinlichkeitsrechnung und mathematischen Statistik erläutert. Eine umfassende Einführung in dieses Thema geben [5] und [7].

### 2.2.1 Wahrscheinlichkeitstheorie

Ein zentraler Begriff der Wahrscheinlichkeitstheorie ist die Zufallsgröße.

**Definition 2.4:** Zufallsgröße  $\mathcal{X}$

Sei  $(\Omega, E, P)$  ein Wahrscheinlichkeitsraum, wobei  $\Omega$  die Menge der Elementarereignisse,  $E$  der Ereignisraum und  $P$  das Wahrscheinlichkeitsmaß ist. Eine Funktion  $\mathcal{X} : \Omega \rightarrow \mathbb{R}$  heißt Zufallsgröße, falls  $\mathcal{X}$  eine messbare Funktion ist, d.h. es gilt

$$\mathcal{X}^{-1}((-\infty, x]) = \{\omega \in \Omega \mid \mathcal{X}(\omega) \leq x\} \in E, \forall x \in \mathbb{R}.$$

Ein fester Wert  $x = \mathcal{X}(\omega)$  heißt Realisierung der Zufallsgröße  $\mathcal{X}$ .

Wichtig für die weiteren Betrachtungen ist die Verteilung einer Zufallsgröße.

**Definition 2.5:** Verteilung einer ZG  $\mathcal{X}$

Die Verteilung einer ZG  $\mathcal{X}$  ist definiert als

$$F(x) = P(\mathcal{X}^{-1}((-\infty, x])) = P(\{\omega \in \Omega \mid \mathcal{X}(\omega) \leq x\});$$

kurz:  $F(x) = P(\mathcal{X} \leq x), \forall x \in \mathbb{R}.$

$F(x)$  ist monoton nichtfallend, rechtsseitig stetig und wegen der Normiertheit des Wahrscheinlichkeitsmaßes  $P$  gilt  $0 \leq F(x) \leq 1, \forall x \in \mathbb{R}$ . Notwendig ist weiterhin die Definition einer stetigen ZG  $\mathcal{X}$ , da in dieser Arbeit nur stetige ZG verwendet werden.

**Definition 2.6:** stetige ZG  $\mathcal{X}$ 

Eine stetige ZG  $\mathcal{X}$  besitzt eine absolut stetige Verteilung  $F(x)$ , d.h. es existiert fast sicher eine Ableitung  $f(x) = F'(x)$ ,  $\forall x \in \mathbb{R}$ . Bezeichnet wird  $f(x)$  als Dichtefunktion. Die Verteilung einer stetigen ZG

$$F(x) = \int_{-\infty}^x f(t) dt, \forall x \in \mathbb{R}$$

mit Wahrscheinlichkeit 1.

**Definition 2.7:** Unabhängigkeit zweier stetiger ZG  $\mathcal{X}$  und  $\mathcal{Y}$ 

Zwei ZG  $\mathcal{X}$  und  $\mathcal{Y}$  mit den jeweiligen Verteilungen  $F_{\mathcal{X}}(x)$  und  $F_{\mathcal{Y}}(y)$  sind unabhängig, wenn

$$F(x, y) = F_{\mathcal{X}}(x) \cdot F_{\mathcal{Y}}(y) .$$

Wichtige Informationen über eine ZG liefern die Momente.

**Definition 2.8:**  $k$  - tes Moment

Sei  $\mathcal{X}$  eine stetige ZG mit Verteilung und  $k \in \mathbb{N}$ . Das  $k$ -te Moment

$$E\mathcal{X}^k = \int_{-\infty}^{\infty} x^k f(x) dx .$$

Voraussetzung für die Existenz der Momente ist die Konvergenz der jeweiligen Integrale. Der Erwartungswert  $\mu$  einer ZG  $\mathcal{X}$  entspricht den ersten Moment, das bedeutet

$$\mu = E\mathcal{X} = \int_{-\infty}^{\infty} x f(x) dx .$$

**Definition 2.9:**  $k$  - tes zentrales Moment

Sei  $\mathcal{X}$  eine stetige ZG mit Verteilung und  $k \in \mathbb{N}$ . Das  $k$ -te zentrale Moment ist

$$D^k \mathcal{X} = E(\mathcal{X} - \mu)^k = \int_{-\infty}^{\infty} (x - \mu)^k f(x) dx .$$

Das zweite zentrale Moment einer ZG  $\mathcal{X}$  ist gleich der Varianz  $\sigma^2$ , d.h.

$$\sigma^2 = D^2 \mathcal{X} = \int_{-\infty}^{\infty} (x - \mu)^2 f(x) dx .$$

Die Varianz  $\sigma^2$  gibt die mittlere quadratische Abweichung vom Erwartungswert  $\mu$  an. Der Erwartungswert  $\mu$  und die Varianz  $\sigma^2$  sind die beiden bekanntesten Momente einer ZG. Weitere Kenngrößen einer ZG  $\mathcal{X}$ , bei der  $\mu$  und  $\sigma^2$  existieren, sind die Standardabweichung

$$\sigma = +\sqrt{\sigma^2}$$

und der Variationskoeffizient

$$\nu = \frac{\sigma}{\mu},$$

mit  $\mu \neq 0$ .

**Definition 2.10:** Schiefe  $\gamma_1$

Die Schiefe einer ZG  $\mathcal{X}$ , wobei das 2. und 3. zentrale Moment existieren, ist definiert als

$$\gamma_1 = \frac{D^3 \mathcal{X}}{\sqrt{(D^2 \mathcal{X})^3}}$$

Gilt für eine ZG  $\mathcal{X}$   $\gamma_1 = 0$ , so ist die Dichtefunktion von  $\mathcal{X}$  symmetrisch bezüglich ihres Erwartungswertes. Ist  $\gamma_1$  kleiner bzw. größer 0, dann heißt die Dichtefunktion linksschief bzw. rechtschief. Veranschaulicht wird dies in Abb. 2.2.

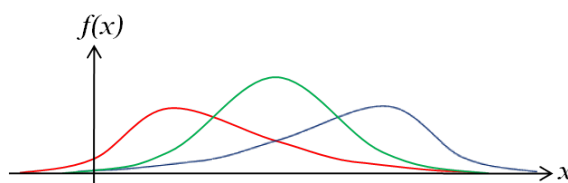


Abb. 2.2: Dichtefunktionen mit  $\gamma_1 = 0$ ,  $\gamma_1 < 0$  und  $\gamma_1 > 0$

**Definition 2.11:** Exzess  $\gamma_2$

Existieren das 2. und 4. zentrale Moment einer ZG  $\mathcal{X}$ , so ist der Exzess, auch Kurtosis oder Wölbung genannt,

$$\gamma_2 = \frac{D^4 \mathcal{X}}{(D^2 \mathcal{X})^2} - 3.$$

Der Exzess  $\gamma_2$  gibt an, ob das Maximum der Dichtefunktion der ZG  $\mathcal{X}$  größer oder kleiner als bei der Dichte der Normalverteilung mit gleicher Varianz ist. Bei der Normalverteilung ist der Exzess 0. Ist  $\gamma_2 > 0$ , so ist das Maximum der Dichtefunktion größer als das der Dichte der zugehörigen Normalverteilung, für  $\gamma_2 < 0$  ist es kleiner (siehe Abb. 2.3).

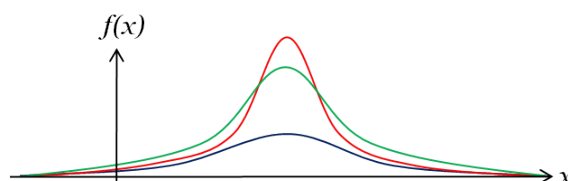


Abb. 2.3: Dichtefunktionen mit  $\gamma_2 = 0$ ,  $\gamma_2 < 0$  und  $\gamma_2 > 0$

Wichtig sind noch die Quantile einer ZG  $\mathcal{X}$ .

**Definition 2.12:**  $p$  - Quantil  $c_p$ 

Sei  $\mathcal{X}$  eine stetige ZG mit einer streng monoton wachsenden Verteilung  $F$ . Dann ist

$$c_p = F^{-1}(p), \quad p \in (0, 1) \quad .$$

Für die Werte  $p = 0.25$ ,  $p = 0.5$  und  $p = 0.75$  haben die Quantile besondere Namen. Ist  $p = 0.25$  bzw.  $p = 0.75$ , so heißen die Quantile unteres bzw. oberes Quartil. Das Quantil  $c_{0.5}$  wird als Median bezeichnet.

**Definition 2.13:** Standardisierte ZG

Eine ZG  $\mathcal{X}$  mit  $\mu = E\mathcal{X} = 0$  und  $\sigma^2 = D^2\mathcal{X} = 1$  heißt standardisiert.

Eine ZG  $\mathcal{X}$ , für die der Erwartungswert  $\mu$  und die Standardabweichung  $\sigma$  existieren, ist durch die folgende lineare Transformation

$$\mathcal{Y} = \frac{\mathcal{X} - \mu}{\sigma}$$

standardisierbar.

Um Aussagen über die Abhängigkeit zwischen zwei ZG treffen zu können, werden folgende Begriffe eingeführt.

**Definition 2.14:** Kovarianz  $cov(\mathcal{X}, \mathcal{Y})$  und Korrelation  $\varrho(\mathcal{X}, \mathcal{Y})$ 

Seien  $\mathcal{X}$  und  $\mathcal{Y}$  zwei ZG, wobei  $D^2\mathcal{X} = \sigma_{\mathcal{X}}^2$  und  $D^2\mathcal{Y} = \sigma_{\mathcal{Y}}^2$  existieren.

$$cov(\mathcal{X}, \mathcal{Y}) = E[(\mathcal{X} - E\mathcal{X})(\mathcal{Y} - E\mathcal{Y})]$$

$$\varrho(\mathcal{X}, \mathcal{Y}) = \frac{cov(\mathcal{X}, \mathcal{Y})}{\sqrt{\sigma_{\mathcal{X}}^2 \cdot \sigma_{\mathcal{Y}}^2}}$$

Die Kovarianz besitzt folgende Eigenschaften:

1.  $cov(\mathcal{X}, \mathcal{Y}) = cov(\mathcal{Y}, \mathcal{X})$
2.  $cov(\mathcal{X}, \mathcal{X}) = D^2\mathcal{X}$
3.  $cov(\mathcal{X}, \mathcal{Y}) = E(\mathcal{X}\mathcal{Y}) - E\mathcal{X} \cdot E\mathcal{Y}$
4.  $cov(\mathcal{X}, \mathcal{Y}) \leq \sqrt{D^2\mathcal{X} \cdot D^2\mathcal{Y}}$

Ist  $cov(\mathcal{X}, \mathcal{Y}) = 0$ , so sind  $\mathcal{X}$  und  $\mathcal{Y}$  unkorreliert. Falls  $cov(\mathcal{X}, \mathcal{Y}) > 0$  ist, heißen  $\mathcal{X}$  und  $\mathcal{Y}$  positiv korreliert. Andernfalls, d.h. wenn  $cov(\mathcal{X}, \mathcal{Y}) < 0$  ist, heißen  $\mathcal{X}$  und  $\mathcal{Y}$  negativ korreliert.

**Bemerkung 2.2:**

Der Korrelationskoeffizient  $\varrho(\mathcal{X}, \mathcal{Y})$  drückt die Art und Stärke des linearen Zusammenhangs zwischen den ZG  $\mathcal{X}$  und  $\mathcal{Y}$  aus. Weiterhin gelten:

1.  $\varrho(\mathcal{X}, \mathcal{Y}) = \varrho(\mathcal{Y}, \mathcal{X})$
2.  $\varrho(\mathcal{X}, \mathcal{Y})$  liegt im Intervall  $[-1, 1]$
3.  $|\varrho(\mathcal{X}, \mathcal{Y})| = 1 \Leftrightarrow \mathcal{Y} = a\mathcal{X} + b$  mit  $a, b \in \mathbb{R}$  und  $a \neq 0$
4.  $\varrho(\mathcal{X}, \mathcal{Y}) = 0 \Leftrightarrow \mathcal{X}$  und  $\mathcal{Y}$  sind unkorreliert
5.  $\mathcal{X}$  und  $\mathcal{Y}$  sind unabhängig  $\Rightarrow \mathcal{X}$  und  $\mathcal{Y}$  sind unkorreliert
6.  $\mathcal{X}, \mathcal{Y}$  unkorreliert und 2 - dim. normalverteilt  $\Rightarrow \mathcal{X}$  und  $\mathcal{Y}$  sind unabhängig

Die ZG  $\mathcal{X}$  und  $\mathcal{Y}$  heißen stark korreliert, wenn  $|\varrho(\mathcal{X}, \mathcal{Y})| \geq 0.8$ . Ist  $|\varrho(\mathcal{X}, \mathcal{Y})| \leq 0.4$ , sind die zwei ZG schwach korreliert.

Die Normalverteilung ist die wichtigste Verteilung in der Wahrscheinlichkeitsrechnung und mathematischen Statistik. Eine ZG  $\mathcal{X}$  ist normalverteilt mit den Realisierungen  $x \in \mathbb{R}$  und den Parametern  $\mu$  und  $\sigma^2$ , wenn sie die Dichte

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

besitzt. Dann schreibt man  $\mathcal{X} \sim N(\mu, \sigma^2)$ . Die Verteilung der Normalverteilung

$$F(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^x e^{-\frac{(t-\mu)^2}{2\sigma^2}} dt .$$

Die Verteilung  $F(x)$  der Normalverteilung ist nicht geschlossen integrierbar, aber standardisierbar. Die Dichtefunktion und Verteilung lauten dann für  $\mathcal{Y} = \frac{\mathcal{X}-\mu}{\sigma}$  mit  $\mathcal{X} \sim N(\mu, \sigma^2)$ :

$$\begin{aligned} \varphi(y) &= \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} \\ \Phi(y) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^y e^{-\frac{t^2}{2}} dt \end{aligned}$$

Weitere Verteilungen werden in A.1.1 vorgestellt. Die Werte der Normalverteilung und der unter A.1.1 vorgestellten liegen u.a. in [3] und [5] tabelliert vor.

### 2.2.2 Deskriptive Statistik

Des Weiteren werden Grundlagen aus der mathematischen Statistik erläutert. Dabei wird nur auf quantitative Merkmale eingegangen. Zudem sei eine konkrete, geordnete Stichprobe aus der Grundgesamtheit des unbekannt verteilten Merkmals  $X$  mit den Merkmalsausprägungen  $x_i$

$(i = 1, \dots, n)$  gegeben, wobei  $n$  der Stichprobenumfang ist. Der arithmetische Mittelwert

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$$

und die empirische Varianz

$$s_X^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 = \frac{1}{n-1} \left( \sum_{i=1}^n x_i^2 - n\bar{x}^2 \right)$$

liefern wichtige Informationen über die Stichprobe. Weitere statistische Kennwerte sind unter den gleichen Gegebenheiten wie hier in A.1.2 erläutert.

Aussagen über den linearen Zusammenhang zwischen zwei quantitativen Merkmalen  $X$  und  $Y$  lassen sich anhand einer Stichprobe mit Hilfe folgender Maßzahlen treffen:

1. empirische Kovarianz:

$$s_{XY} = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) = \frac{1}{n-1} \left( \sum_{i=1}^n x_i y_i - n\bar{x}\bar{y} \right)$$

2. empirischer Korrelationskoeffizient nach Bravais - Pearson:

$$r_{XY} = \frac{s_{XY}}{s_X s_Y} = \frac{\sum_{i=1}^n x_i y_i - n\bar{x}\bar{y}}{\sqrt{\left( \sum_{i=1}^n x_i^2 - n\bar{x}^2 \right) \left( \sum_{i=1}^n y_i^2 - n\bar{y}^2 \right)}}$$

Für die empirische Kovarianz und Korrelation gelten die gleichen Beziehungen und Eigenschaften wie bei Definition 2.14 und Bemerkung 2.2.

### 2.2.3 Multiple lineare Regression

Als Letztes in diesem Kapitel wird die multiple lineare Regression betrachtet. Aufbauend auf die einfache lineare Regression, vgl. [5], wird bei der multiplen Regression ein funktionaler Zusammenhang der Form

$$y(x_1, \dots, x_k) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k = \sum_{j=0}^k \beta_j x_j \quad (2.1)$$

zwischen den  $k+1$  ( $k > 1$ ) quantitativen Merkmalen  $Y, X_1, X_2, \dots, X_k$  einer gegebenen Stichprobe vom Umfang  $n$  gesucht. Hierbei sei zur leichteren Schreibweise  $x_0 = 1$ . Die Gleichung (2.1) wird multiple Regressionsgleichung vom Regressanden  $Y$  (Zielgröße) auf die Regressoren  $X_1, X_2, \dots, X_k$  (Einflussgrößen) genannt. Grundlage der multiplen Regressionsgleichung ist das Modell

$$\mathcal{Y} = \beta_0 + \beta_1 \mathcal{X}_1 + \beta_2 \mathcal{X}_2 + \dots + \beta_k \mathcal{X}_k + \mathcal{U},$$

wobei die  $\beta_j$  die theoretischen Regressionskoeffizienten sind und  $\mathcal{U}$  eine zufällige Störgröße ist. Vorausgesetzt ist hierbei, dass die Regressoren  $X_j$  unabhängig sind und  $\mathcal{U} \sim N(0, \sigma^2)$  ist. Die  $\beta_j$ ,  $j = 0, \dots, k$  sind unbekannt und müssen geschätzt werden, um eine konkrete Regressionsgleichung

$$\hat{y}(x_1, \dots, x_k) = b_0 + b_1 x_1 + b_2 x_2 + \dots + b_k x_k = \sum_{j=0}^k b_j x_j \quad (2.2)$$

zu erhalten. Die Schätzung der Regressionskoeffizienten erfolgt nach der Methode der kleinsten Quadrate, d.h. die Quadratsumme der Residuen  $u_i$  wird minimiert.

$$S(\mathbf{b}) = \sum_{i=1}^n u_i^2 = \sum_{i=1}^n (y_i - \hat{y}_i)^2 = \sum_{i=1}^n (y_i - b_0 - b_1 x_{i1} - b_2 x_{i2} - \dots - b_k x_{ik})^2 \rightarrow \min_{b_j} \quad (2.3)$$

In Matrixschreibweise lautet

$$S(\mathbf{b}) = \mathbf{u}^T \mathbf{u} = (\mathbf{y} - \hat{\mathbf{y}})^T (\mathbf{y} - \hat{\mathbf{y}}) = (\mathbf{y} - \mathbf{X}\mathbf{b})^T (\mathbf{y} - \mathbf{X}\mathbf{b}) \rightarrow \min_{\mathbf{b}}, \quad (2.4)$$

wobei

$$\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} \in \mathbb{R}^n, \quad \mathbf{b} = \begin{pmatrix} b_0 \\ b_1 \\ \vdots \\ b_k \end{pmatrix} \in \mathbb{R}^{k+1} \text{ und } \mathbf{X} = \begin{pmatrix} 1 & x_{11} & \cdots & x_{1k} \\ 1 & x_{21} & \cdots & x_{2k} \\ \vdots & \vdots & & \vdots \\ 1 & x_{n1} & \cdots & x_{nk} \end{pmatrix} \in \mathbb{R}^{n \times (k+1)}.$$

Die notwendige Bedingung für das Minimum ist, dass die partiellen Ableitungen  $\frac{\partial S}{\partial b_j}$  gleich 0 werden, für  $j = 0, \dots, k$ . Zunächst wird (2.4) umgeformt, dann partiell differenziert und 0 gesetzt.

$$\begin{aligned} S(\mathbf{b}) &= (\mathbf{y} - \mathbf{X}\mathbf{b})^T (\mathbf{y} - \mathbf{X}\mathbf{b}) \\ &= \mathbf{y}^T \mathbf{y} - 2\mathbf{b}^T \mathbf{X}^T \mathbf{y} + \mathbf{b}^T \mathbf{X}^T \mathbf{X} \mathbf{b} \rightarrow \min_{\mathbf{b}} \\ \frac{\partial S}{\partial \mathbf{b}} &= -2\mathbf{X}^T \mathbf{y} + 2\mathbf{X}^T \mathbf{X} \mathbf{b} \stackrel{!}{=} \mathbf{0} \end{aligned}$$

Dadurch ergibt sich das lineare Gleichungssystem der Gauß'schen Normalgleichungen:

$$\mathbf{X}^T \mathbf{X} \mathbf{b} = \mathbf{X}^T \mathbf{y} \quad (2.5)$$

Aus (2.5) folgt, dass

$$\mathbf{X}^T \mathbf{u} = \mathbf{X}^T (\mathbf{y} - \mathbf{X}\mathbf{b}) = \mathbf{0}$$

und wiederum daraus, dass

$$\sum_{i=1}^n u_i = 0 \quad (2.6)$$

ist. Falls  $\mathbf{X}^T \mathbf{X}$  regulär ist, kann  $\mathbf{b}$  aus

$$\mathbf{b} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y} \quad (2.7)$$

berechnet werden. Die hinreichende Bedingung für das Minimum ist genau dann erfüllt, wenn die Matrix der zweiten partiellen Ableitungen an der Extremwertstelle positiv definit ist. Numerisch günstiger als die Berechnung von  $\mathbf{b}$  über 2.7 ist die direkte Lösung von (2.5), z.B. mittels Cholesky-Zerlegung (siehe [6]), da  $\mathbf{X}^T \mathbf{X}$  symmetrisch ist. Sei

$$\mathbf{C} = \begin{pmatrix} s_{X_1}^2 & s_{X_1 X_2} & \cdots & s_{X_1 X_k} \\ s_{X_2 X_1} & s_{X_2}^2 & \cdots & s_{X_2 X_k} \\ \vdots & \vdots & \ddots & \vdots \\ s_{X_k X_1} & s_{X_k X_2} & \cdots & s_{X_k}^2 \end{pmatrix} \in \mathbb{R}^{k \times k}$$

die (empirische) Kovarianzmatrix. Die (empirische) Korrelationsmatrix  $\mathbf{R}$  ergibt sich mit Hilfe von  $\mathbf{C}$  zu

$$\mathbf{R} = \begin{pmatrix} \frac{s_{X_1 X_1}}{s_{X_1} s_{X_1}} & \frac{s_{X_1 X_2}}{s_{X_1} s_{X_2}} & \cdots & \frac{s_{X_1 X_k}}{s_{X_1} s_{X_k}} \\ \frac{s_{X_2 X_1}}{s_{X_2} s_{X_1}} & \frac{s_{X_2 X_2}}{s_{X_2} s_{X_2}} & \cdots & \frac{s_{X_2 X_k}}{s_{X_2} s_{X_k}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{s_{X_k X_1}}{s_{X_k} s_{X_1}} & \frac{s_{X_k X_2}}{s_{X_k} s_{X_2}} & \cdots & \frac{s_{X_k X_k}}{s_{X_k} s_{X_k}} \end{pmatrix} = \begin{pmatrix} 1 & r_{X_1 X_2} & \cdots & r_{X_1 X_k} \\ r_{X_2 X_1} & 1 & \cdots & r_{X_2 X_k} \\ \vdots & \vdots & \ddots & \vdots \\ r_{X_k X_1} & r_{X_k X_2} & \cdots & 1 \end{pmatrix}. \quad (2.8)$$

Im Folgenden werden Aussagen über die Güte der geschätzten multiplen Regressionsgleichung getroffen. Dazu sei

$$\mathbf{r}_{YX} = \begin{pmatrix} r_{YX_1} \\ r_{YX_2} \\ \vdots \\ r_{YX_k} \end{pmatrix} \in \mathbb{R}^k$$

der Vektor der einfachen Korrelationen von  $Y$  mit dem jeweiligen Merkmal  $X_j$  ( $j = 1, \dots, k$ ). Das Quadrat des multiplen linearen Korrelationskoeffizienten berechnet sich aus

$$r_{YX_1 X_2 \dots X_k}^2 = \mathbf{r}_{YX}^T \mathbf{R}^{-1} \mathbf{r}_{YX}. \quad (2.9)$$

**Definition 2.15:** Streuungsmaße

Sei eine multiple Regression eines Merkmals  $Y$  auf  $k$  Merkmale  $X_j$  gegeben. Dann lässt sich die Gesamtvarianz

$$s_Y^2 = \frac{1}{n-1} \sum_{i=1}^n (y_i - \bar{y})^2,$$

die erklärte Varianz

$$s_{\hat{Y}}^2 = \frac{1}{n-1} \sum_{i=1}^n (\hat{y}_i - \bar{y})^2$$

und die Restvarianz

$$s^2 = \frac{1}{n-k-1} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

definieren.



**Satz 2.1:** Zerlegung der Gesamtvarianz

Gegeben sei eine multiple Regression eines Merkmals  $Y$  auf  $k$  Merkmale  $X_j$ , wobei die Schätzung der multiplen Regressionsgleichung nach der Methode der kleinsten Quadrate erfolgte. Dann gilt

$$(n-1)s_Y^2 = (n-k-1)s^2 + (n-1)s_{\hat{Y}}^2.$$

**Beweis:**

Sei  $\mathbf{y}_m = \begin{pmatrix} \bar{y} \\ \vdots \\ \bar{y} \end{pmatrix} \in \mathbb{R}^n$ . Die Streuungsmaße lassen sich dadurch schreiben als:

$$\begin{aligned} (n-1)s_Y^2 &= (\mathbf{y} - \mathbf{y}_m)^T (\mathbf{y} - \mathbf{y}_m) \\ (n-1)s_{\hat{Y}}^2 &= (\mathbf{X}\mathbf{b} - \mathbf{y}_m)^T (\mathbf{X}\mathbf{b} - \mathbf{y}_m) \\ (n-k-1)s^2 &= (\mathbf{y} - \mathbf{X}\mathbf{b})^T (\mathbf{y} - \mathbf{X}\mathbf{b}) \end{aligned}$$

Durch Umformung des  $(n-1)$ -fachen der Gesamtvarianz erhält man

$$\begin{aligned} (n-1)s_Y^2 &= (\mathbf{y} - \mathbf{y}_m)^T (\mathbf{y} - \mathbf{y}_m) \\ &= (\mathbf{y} - \mathbf{X}\mathbf{b} + \mathbf{X}\mathbf{b} - \mathbf{y}_m)^T (\mathbf{y} - \mathbf{X}\mathbf{b} + \mathbf{X}\mathbf{b} - \mathbf{y}_m) \\ &= ((\mathbf{y} - \mathbf{X}\mathbf{b}) + (\mathbf{X}\mathbf{b} - \mathbf{y}_m))^T ((\mathbf{y} - \mathbf{X}\mathbf{b}) + (\mathbf{X}\mathbf{b} - \mathbf{y}_m)) \\ &= (\mathbf{y} - \mathbf{X}\mathbf{b})^T (\mathbf{y} - \mathbf{X}\mathbf{b}) + (\mathbf{y} - \mathbf{X}\mathbf{b})^T (\mathbf{X}\mathbf{b} - \mathbf{y}_m) + (\mathbf{X}\mathbf{b} - \mathbf{y}_m)^T (\mathbf{y} - \mathbf{X}\mathbf{b}) \\ &\quad + (\mathbf{X}\mathbf{b} - \mathbf{y}_m)^T (\mathbf{X}\mathbf{b} - \mathbf{y}_m) \\ &= (\mathbf{y} - \mathbf{X}\mathbf{b})^T (\mathbf{y} - \mathbf{X}\mathbf{b}) + 2(\mathbf{y} - \mathbf{X}\mathbf{b})^T (\mathbf{X}\mathbf{b} - \mathbf{y}_m) + (\mathbf{X}\mathbf{b} - \mathbf{y}_m)^T (\mathbf{X}\mathbf{b} - \mathbf{y}_m) \\ &= (n-k-1)s^2 + 2(\mathbf{y} - \mathbf{X}\mathbf{b})^T (\mathbf{X}\mathbf{b} - \mathbf{y}_m) + (n-1)s_{\hat{Y}}^2 \end{aligned}$$

Jetzt bleibt noch zu zeigen, dass  $(\mathbf{y} - \mathbf{X}\mathbf{b})^T (\mathbf{X}\mathbf{b} - \mathbf{y}_m) = 0$  ist.

$$\begin{aligned} (\mathbf{y} - \mathbf{X}\mathbf{b})^T (\mathbf{X}\mathbf{b} - \mathbf{y}_m) &= \mathbf{y}^T \mathbf{X}\mathbf{b} - \mathbf{y}^T \mathbf{y}_m - \mathbf{b}^T \mathbf{X}^T \mathbf{X}\mathbf{b} + \mathbf{b}^T \mathbf{X}^T \mathbf{y}_m \\ &= \mathbf{b}^T (\mathbf{X}^T \mathbf{y} - \mathbf{X}^T \mathbf{X}\mathbf{b}) - (\mathbf{y} - \mathbf{X}\mathbf{b})^T \mathbf{y}_m \end{aligned}$$

Aufgrund von (2.5) ist  $\mathbf{X}^T \mathbf{y} - \mathbf{X}^T \mathbf{X}\mathbf{b}$  identisch mit dem Nullvektor. Mit (2.6) folgt

$$\begin{aligned} (\mathbf{y} - \mathbf{X}\mathbf{b})^T (\mathbf{X}\mathbf{b} - \mathbf{y}_m) &= -(\mathbf{y} - \mathbf{X}\mathbf{b})^T \mathbf{y}_m \\ &= -\sum_{i=1}^n (y_i - \hat{y}_i) \bar{y} \\ &= -\bar{y} \sum_{i=1}^n u_i \\ &= 0 \end{aligned}$$

und schließlich die Gültigkeit des Satzes.

□

Ein Maß für die Güte der multiplen Regression ist das Verhältnis aus erklärter Varianz und Gesamtvarianz.

**Definition 2.16:** Bestimmtheitsmaß  $b_{YX_1 \dots X_k}$  und Unbestimmtheitsmaß  $u_{YX_1 \dots X_k}$

Gegeben sei eine multiple Regression und die Gültigkeit des Satzes 2.1. Das Bestimmtheitsmaß

$$b_{YX_1 \dots X_k} = \frac{s_{\hat{Y}}^2}{s_Y^2}$$

und das Unbestimmtheitsmaß

$$u_{YX_1 \dots X_k} = \frac{(n - k - 1) s^2}{(n - 1) s_Y^2}.$$

Durch diese Definition ergeben sich folgende Eigenschaften von  $b_{YX_1 \dots X_k}$  und  $u_{YX_1 \dots X_k}$ :

1.  $b_{YX_1 \dots X_k} + u_{YX_1 \dots X_k} = 1$
2.  $0 \leq b_{YX_1 \dots X_k} \leq 1, \quad 0 \leq u_{YX_1 \dots X_k} \leq 1$
3.  $s^2 = \frac{n - 1}{n - k - 1} s_Y^2 (1 - b_{YX_1 \dots X_k}) = \frac{n - 1}{n - k - 1} s_Y^2 u_{YX_1 \dots X_k}$

Das Bestimmtheitsmaß  $b_{YX_1 \dots X_k}$  gibt den Anteil der Varianz von Y an, der durch die Merkmale  $X_1, X_2, \dots, X_k$  erklärt wird. Somit ist  $b_{YX_1 \dots X_k}$  ein Maß dafür, wie gut die Zielgröße durch die Einflussgrößen erklärt wird. Weiterhin entspricht

$$b_{YX_1 \dots X_k} = r_{YX_1 \dots X_k}^2$$

bei der multiplen linearen Regression (vgl. [15]).

## 3 Sensitivitätsanalyse

Die Sensitivitätsanalyse dient zur Ermittlung der einflussreichsten Parameter zu unterschiedlichen Zeitpunkten im Lösungsprozess einer MOA. Im Optimierungsprozesses ist vor allem von Interesse, dass nur die Parameter in die MOA integriert sind, die relevant für die Optimierung sind. Bei der a-priori Analyse werden zu Beginn der Optimierung die einflussreichsten Parameter ermittelt, wodurch sichergestellt werden kann, dass nur die wichtigsten Parameter im Optimierungsprozess verwendet werden. Außerdem ist es möglich, während des Lösungsprozesses die sensitiven Parameter zu ermitteln, wodurch gegebenenfalls mit einer reduzierten Parametermenge weiter gerechnet werden kann. Nach erfolgter Optimierung dient die Sensitivitätsanalyse zur Überprüfung der vorhergehenden Analysen, falls aus Gründen der Genauigkeit oder anderer keine Parameter aus der MOA entfernt wurden und um die Zusammenhänge in der Optimierungsaufgabe zu analysieren. Dabei sind auch die technischen und physikalischen Eigenschaften der MOA zu beachten, um scheinbar relevante Parametereinflüsse zu vermeiden. Die Kenntnis der sensitiven Parameter ist auch für eine nachfolgende Robustheitsbewertung von Bedeutung.

Als Erstes wird die Analyse der Sensitivität vor dem Beginn der Lösung einer MOA betrachtet. Danach werden anhand der erhaltenen Lösungen aus der Optimierung Aussagen über die Sensitivität der Parameter getroffen. Die a-posteriori Sensitivität kann auch während des Optimierungsprozesses durchgeführt werden, wenn für die statistischen Analyseverfahren genügend Lösungen zur Auswertung bereit stehen.

### 3.1 A-priori Sensitivität

Ist vor Beginn der Optimierung einer MOA bereits von Interesse, welche Parameter den größten Einfluss auf die Zielgrößen haben, so müssen die Zielfunktionen in Abhängigkeiten von ausgewählten Parameterwerten untersucht werden. Gründe für eine a-priori Analyse sind beispielsweise die Reduzierung einer großen Anzahl an Parametern auf die Wesentlichen oder das Ausmachen von Parametern, die mit erhöhter Genauigkeit bestimmt werden müssen, um präzise Zielfunktionswerte zu erhalten. Einerseits steht hier das Verringern der Rechenzeit des Optimierungsverfahren im Vordergrund, andererseits aber auch die Exaktheit der ermittelten Lösungen. Zudem sollte die Untersuchung im Voraus nicht zu viel Rechenzeit in Anspruch nehmen.

Die Berechnung von Zielfunktionswerten der hier betrachteten MOA's geschieht meistens durch computergestützte Simulationen. Das ist bedingt durch die praxisbezogenen Problemstellungen. Dadurch steht kein exakter analytischer Ausdruck der Zielfunktionen zur Verfügung. Zur

Betrachtung von Sensitivitäten der Parameter werden numerische Methoden verwendet. Sei  $\mathbf{x} = (x_1, x_2, \dots, x_n)^T$  der Vektor der Ausgangswerte der Parameter. Der Vektor  $\mathbf{x}$  ist in den meisten praxisrelevanten Fällen durch den Ausgangszustand des zu optimierenden Problems gegeben. Die daraus resultierenden Zielfunktionswerte sind dann ebenfalls gegeben. Andernfalls kann  $\mathbf{x}$  sinnvoll aus den Restriktionsbereich gewählt werden. In  $\mathbf{x}$  wird schließlich der Gradient  $\nabla f(\mathbf{x})$  diskretisiert. Das geschieht durch Annäherung der partiellen Ableitungen über Differenzenformeln. Beispiele für Differenzenformeln sind der vorwärtige

$$D_+ f(\mathbf{x}) = \frac{f(\mathbf{x} + h) - f(\mathbf{x})}{h},$$

der rückwärtige

$$D_- f(\mathbf{x}) = \frac{f(\mathbf{x}) - f(\mathbf{x} - h)}{h}$$

oder der zentrale Differenzenquotient

$$D_z f(\mathbf{x}) = \frac{f(\mathbf{x} + h) - f(\mathbf{x} - h)}{2h}.$$

Unter Nutzung der zentralen Differenzenformel ist der diskretisierte Gradient

$$\nabla_D f(\mathbf{x}) = \begin{pmatrix} \frac{f(\mathbf{x}_1 + h) - f(\mathbf{x}_1 - h)}{2h} \\ \vdots \\ \frac{f(\mathbf{x}_n + h) - f(\mathbf{x}_n - h)}{2h} \end{pmatrix}.$$

Der Vorteil des zentralen Differenzenquotienten liegt in der höheren Approximationsordnung (vgl. [14]). Jedoch ist zu prüfen, ob der Genauigkeitsgewinn gerechtfertigt ist, da im Vergleich zu den anderen die doppelte Anzahl an Zielfunktionsauswertungen zur Berechnung notwendig ist. Zusätzlich zu den berechneten Gradienten im Punkt  $\mathbf{x}$  können auch in anderen Punkten aus dem Restriktionsbereich der gegebenen MOA Gradienten gebildet werden. Dadurch steigt jedoch erheblich der Rechenaufwand, da die Anzahl der Zielfunktionsauswertungen mit der Anzahl der zu berechnenden Gradienten wächst. Im Allgemeinen wird sich  $\nabla_D f(\mathbf{x})$  von den Gradienten in anderen Punkten mehr oder weniger stark unterscheiden. Die betragsmäßig größten Komponenten (Anstiege) der betrachteten Gradienten sind Anzeichen dafür, dass der zur Komponente gehörende Parameter sensitiv ist. Ausschlaggebend für einflussreiche Parameter ist aber der Beitrag, der zum Zielfunktionswert geleistet wird. Daher wird das Skalarprodukt

$$\nabla_D f(\mathbf{x})^T \Delta \mathbf{x}, \tag{3.1}$$

betrachtet, wobei  $\Delta \mathbf{x} \in \mathbb{R}^n$  ein Vektor aus Änderungen  $\Delta x_i$  der Parameterwerte ist. Sei

$$\Delta \mathbf{x}_j = (\delta_{ij} \Delta x_i),$$

für  $i = 1, \dots, n$ . Dann beschreibt  $\nabla_D f(\mathbf{x})^T \Delta \mathbf{x}_j$  annähernd die Veränderung des Zielfunktionswertes bei Änderung des  $j$ -ten Parameters um  $\Delta x_j$ . Besonders bei stark nichtlinearen Zielfunktionen

muss  $\Delta x_j$  klein gewählt werden, um relativ genaue Ergebnisse zu erhalten. Ist aus technischen und physikalischen Eigenschaften der Optimierungsaufgabe oder aus Erfahrung mit ähnlichen Problemstellungen bekannt, dass zwischen dem  $j$ -ten Parameter und der Zielgröße ein linearer bzw. fast linearer Zusammenhang besteht, so liefert die Berechnung mit (3.1) auch für größere  $\Delta x_j$  gute Ergebnisse. Die Wahl der  $\Delta x_j$  kann beispielsweise durch  $p_j \cdot l_j$  mit  $p_j \in [-1, 1]$  für  $j = 1, \dots, n$  erfolgen, wobei  $l_j$  die Länge des  $j$ -ten Parameterbereiches ist. Zum besseren Vergleich der Einflüsse der einzelnen Parameter auf die Zielgröße sollte  $p_j = p = \text{konstant}$ ,  $\forall j$ , gewählt werden. Die Parameter können anschließend anhand der Größe der Zielfunktionswertänderungen im jeweiligen Punkt  $\mathbf{x}$  sortiert werden. Das spiegelt jedoch nur den Einfluss der Parameter in der Umgebung von  $\mathbf{x}$  wieder. Um globale Entscheidungen treffen zu können, müsste die Änderung der Zielfunktionswerte in mehreren Punkten des Restriktionsbereichs betrachtet und verglichen werden. Abschließend können die Resultate graphisch, z.B. in Balkendiagrammen, dargestellt werden. Zur Veranschaulichung wird eine a-priori Analyse der folgenden einkriteriellen Optimierungsaufgabe mit einer modifizierten Zielfunktion nach Branin (vgl. [16]) betrachtet.

**Beispiel 1:**

$$\begin{aligned} f(x, y) &= \left( y - \left( \frac{x}{\pi} \right)^2 + \frac{5x}{\pi} - 6 \right)^2 + 10 \cos x + 10 \rightarrow \min \\ |x - 9,5| - 3,5 &\leq 0 \\ |y| - 5 &\leq 0 \\ \mathbf{x} &= \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 \end{aligned}$$

Der Graph der Zielfunktion in dem durch Ungleichungen beschriebenen Restriktionsbereich  $R$  ist in 3.1 dargestellt. Eine größere Ansicht befindet sich unter 3.2. Das Minimum befindet sich in  $\mathbf{x}^* = \begin{pmatrix} 3\pi \\ 0 \end{pmatrix}$  mit  $f(\mathbf{x}^*) = 0$ . Ausgangspunkt der Analyse sei  $\mathbf{x}^{(1)} = \begin{pmatrix} \frac{5\pi}{2} \\ 2 \end{pmatrix}$  mit den zugehörigen

Gradienten  $\nabla f(\mathbf{x}^{(1)}) = \begin{pmatrix} -10 \\ \frac{9}{2} \end{pmatrix}$  und den Zielfunktionswert  $f(\mathbf{x}^{(1)}) = 15,0625$ . Alle Gradienten wurden hierbei analytisch berechnet. Entsprechend eines 10%-igen Anteils am jeweiligen Parameterintervall werden  $\Delta x^{(1)} = 0,7$  und  $\Delta y^{(1)} = -1$  gesetzt. Die Vorzeichen der Änderungen werden derart gewählt, dass die Zielfunktionswertänderung in Optimierungsrichtung geschieht.

$$\nabla f(\mathbf{x}^{(1)})^T \Delta \mathbf{x}^{(1)} = \begin{pmatrix} -10 & \frac{9}{2} \end{pmatrix} \begin{pmatrix} 0,7 \\ -1 \end{pmatrix} = -7 - 4,5 = -11,5$$

Daraus folgt, dass bei Erhöhung von  $x$  um 0,7 der Zielfunktionswert um 7 Einheiten fällt und bei Verringerung von  $y$  um 1 der Zielfunktionswert um 4,5 Einheiten sinkt. Werden beide Änderungen zugleich berücksichtigt, nimmt der Zielfunktionswert um 11,5 Einheiten ab. Das gilt natürlich nur im Bezug auf  $\mathbf{x}^{(1)}$ . Die Näherung für  $f(\mathbf{x}^{(1)} + \Delta \mathbf{x}^{(1)})$  ergibt sich zu 3,5625 (exakt: 4,9987). Des Weiteren wird noch der Punkt  $\mathbf{x}^{(2)} = \begin{pmatrix} \frac{7\pi}{2} \\ -2 \end{pmatrix}$  betrachtet.

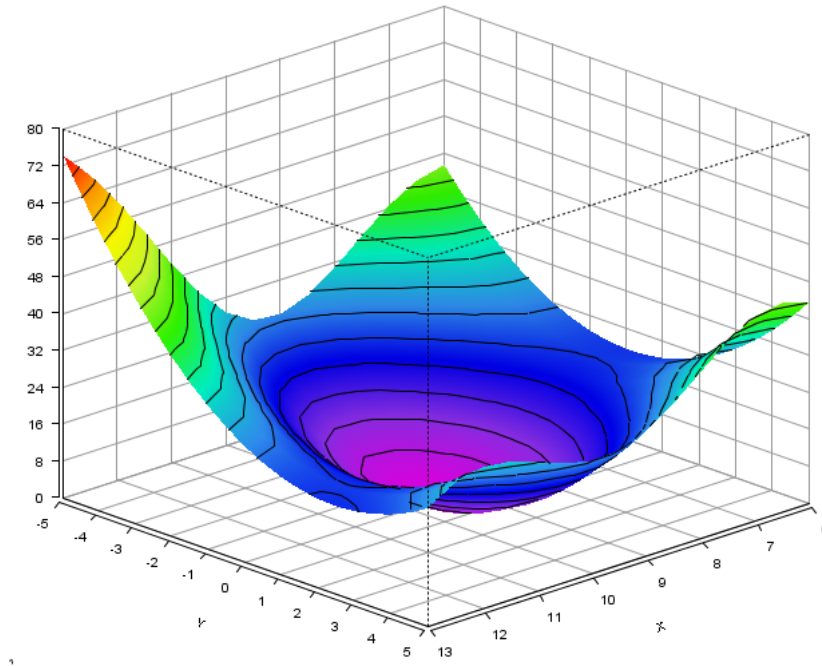


Abb. 3.1: Modifizierte Branin-Funktion im Bereich  $[6, 13] \times [-5, 5]$

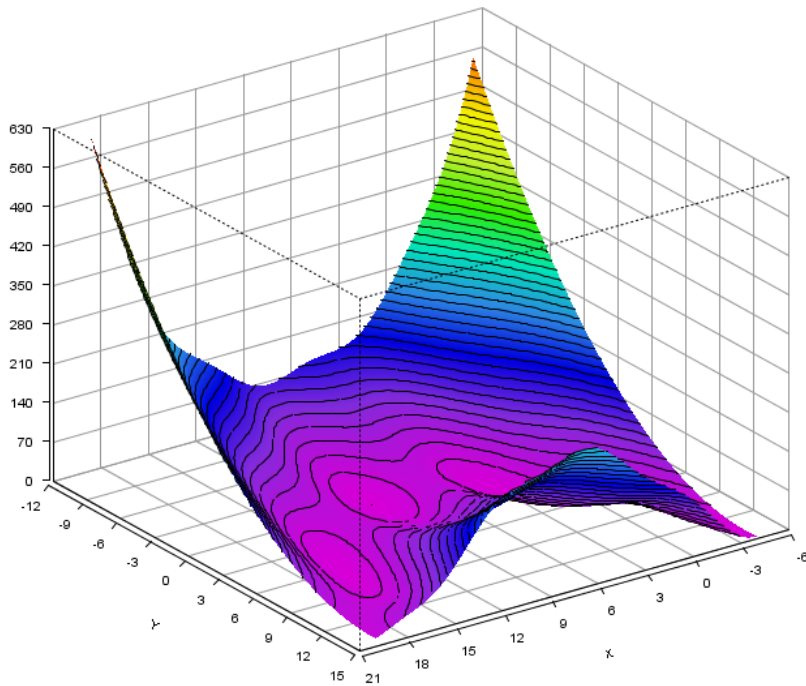


Abb. 3.2: Modifizierte Branin-Funktion im Bereich  $[-6, 21] \times [-12, 15]$

Dabei ist  $\nabla f(\mathbf{x}^{(2)}) = \begin{pmatrix} 10 - \frac{11}{\pi} \\ -\frac{11}{2} \end{pmatrix}$ ,  $\Delta \mathbf{x}^{(2)} = \begin{pmatrix} -10 \\ \frac{9}{2} \end{pmatrix}$  und schließlich

$$\nabla f(\mathbf{x}^{(2)})^T \Delta \mathbf{x}^{(2)} = \left( 10 - \frac{11}{\pi} \quad -\frac{11}{2} \right) \begin{pmatrix} -10 \\ \frac{9}{2} \end{pmatrix} = -7 + \frac{77}{10\pi} - 5,5 = -12,5 + \frac{77}{10\pi} \approx -10,0490.$$

Der Zielfunktionswert  $f(\mathbf{x}^{(2)} + \Delta \mathbf{x}^{(2)})$  ist 5,0135 (exakt: 5,3912). Daraus lässt sich folgern, dass beide Parameter zur Beschreibung des Zielfunktionswertes notwendig sind, da sie entscheidende Beiträge zum Zielfunktionswert leisten. Die Sensitivität von  $x$  ist dabei höher einzuschätzen als die von  $y$ . Das wird auch durch den analytischen Ausdruck der Zielfunktion und der Abb. 3.1 verdeutlicht. In der Abb. 3.3 sind die Ergebnisse veranschaulicht.

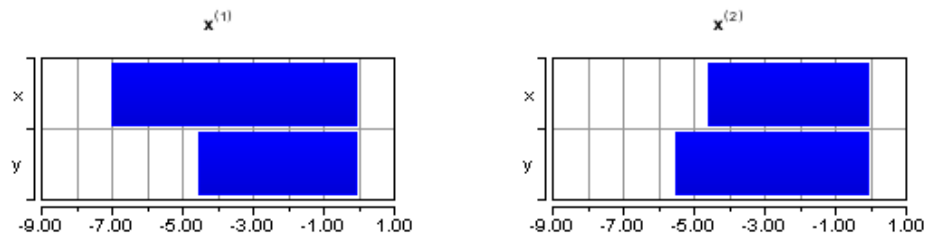


Abb. 3.3: Balkendiagramme der Beiträge zum Zielfunktionswert in  $\mathbf{x}^{(1)}$  und  $\mathbf{x}^{(2)}$

## 3.2 A-posteriori Sensitivität

In diesem Abschnitt wird die Sensitivitätsanalyse im Anschluss an eine erfolgreiche Optimierung einer MOA betrachtet. Wie bereits im vorherigen Abschnitt stehen keine expliziten Ausdrücke der Zielfunktionen zur Verfügung. Die Analyse beruht im Wesentlichen auf der Auswertung der Daten der Paretomenge und Paretofront. Diese Daten werden als eine Stichprobe vom Umfang  $n$  gedeutet, wobei  $n$  die Anzahl der erhaltenen Lösungen ist. Unter der Grundgesamtheit, aus der die Stichprobe erfolgt, wird die Menge aller Lösungen, im Allgemeinen aller paretooptimalen Lösungen, verstanden. Geeignete Lösungsverfahren, um eine Stichprobe zu generieren, sind beispielsweise Genetische Algorithmen, da diese eine gut ausgebildete Paretomenge liefern. Im Kapitel 5 werden schließlich die im Folgenden vorgestellten Sensitivitätsmaße anhand eines Beispiels aus der Praxis diskutiert und verglichen.

### 3.2.1 Regression und Korrelation

Aufgrund der Zusammenhänge zur Statistik, werden die Parameter und Zielgrößen im Folgenden groß geschrieben. Um Aussagen über die Sensitivität der Parameter  $X_j$  bzgl. der Zielgrößen  $Y_i$  treffen zu können, wird für jedes  $Y_i$  ein eigenes multiples Regressionsmodell zu Grunde gelegt. Vorerst wird die Analyse für eine Zielgröße  $Y_i = Y$  durchgeführt. Die dabei verwendeten Methoden sind analog für die anderen Zielgrößen durchführbar. Hierbei seien  $Y$  und  $X_1$  bis  $X_k$

quantitative, stetige Merkmale. Bestimmt wird zuerst ein funktionaler Zusammenhang der Form

$$\hat{y}(x_1, \dots, x_k) = b_0 + b_1x_1 + b_2x_2 + \dots + b_kx_k = \sum_{j=0}^k b_jx_j ,$$

wobei die  $b_j$  aus den IGIS der Gauß'schen Normalgleichungen bestimmt werden. Für die weiteren Untersuchungen ist es notwendig, dass die multiple Regression gut bestimmt ist, da sonst die Folgerungen für die Sensitivität an Aussagekraft verlieren. In der Regel ist das erfüllt, wenn das Bestimmtheitsmaß  $b_{YX_1 \dots X_k} \geq 0.8$  ist. Sollte das nicht der Fall sein, muss die Regressionsgleichung um einige, gut ausgewählte Glieder erweitert werden. Hat beispielsweise der  $j$ -te Parameter quadratischen Einfluss auf  $Y$ , so kann die Regression um  $b_{k+1}x_{k+1} = b_{k+1}x_j^2$  erweitert werden. Dies lässt sich gut aus der 2D-Relationsgraphik im Modul MCA der ET ablesen (siehe Abb. 3.4). Aus der in [5] erläuterten Residualanalyse lassen sich ähnliche Aussagen generieren.

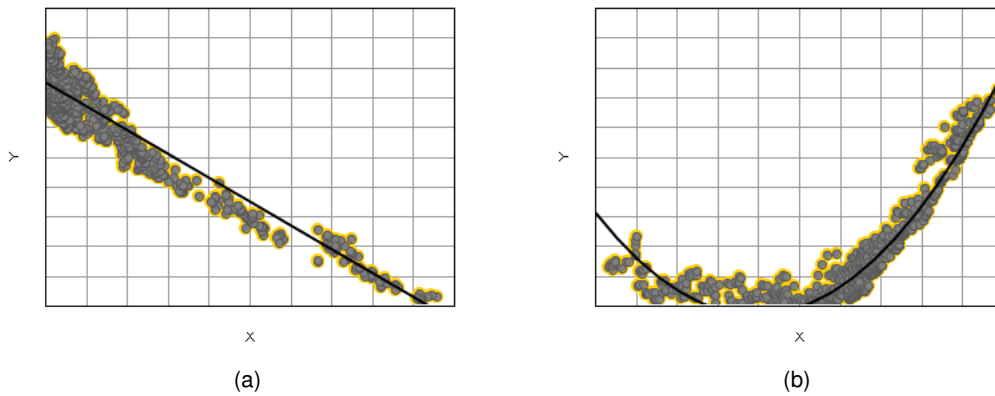


Abb. 3.4: (a) starker linearer und (b) starker quadratischer Zusammenhang zwischen  $Y$  und  $X$

Ziel hierbei ist nicht die Erhöhung des Bestimmtheitsmaßes durch die Hinzunahme einer Vielzahl von Regressoren, da stets

$$b_{YX_1 \dots X_k} \leq b_{YX_1 \dots X_k X_{k+1}} \quad (3.2)$$

gilt, sondern die Erklärung von  $Y$  durch wenige, gut angepasste Parameter. Zusätzlich können die Regressionskoeffizienten  $b_j$  darauf geprüft werden, ob sie signifikant von 0 verschieden sind. Dabei kann die Nullhypothese

$$H_0 : \beta_j = 0 \quad \text{gegen die Alternative} \quad H_1 : \beta_j \neq 0 .$$

getestet werden. Gilt für ein vorgegebenes Niveau  $p$

$$\frac{|b_j|}{s_{b_j}} > c_{t(n-k-1); 1-\frac{p}{2}} ,$$

so ist die Nullhypothese zu verwerfen (vgl. [13]). Hierbei ist  $c_{t(n-k-1); 1-\frac{p}{2}}$  das  $(1 - \frac{p}{2})$  - Quantil der Student'schen  $t$ -Verteilung mit  $n - k - 1$  Freiheitsgraden und  $s_{b_j}$  die Standardabweichung



des Regressionskoeffizienten  $b_j$ . Wird  $H_0$  nicht abgelehnt, so kann die multiple Regressionsgleichung um das  $j$ -te Glied reduziert werden, ohne das Bestimmtheitsmaß wesentlich zu reduzieren. Generell sollte der Stichprobenumfang  $n$  viel größer als die Anzahl der Regressoren  $k'$  der u.U. veränderten Regressionsgleichung sein. Ist das nicht der Fall, kann das korrigierte Bestimmtheitsmaß (vgl. [9])

$$\tilde{b}_{YX_1 \dots X_{k'}} = 1 - \frac{n-1}{n-k'-1} (1 - b_{YX_1 \dots X_{k'}}) \quad (3.3)$$

verwendet werden. Es ist auch besser zum Vergleich mehrerer unterschiedlicher Regressionsmodelle geeignet, da die jeweilige Anzahl von Regressoren berücksichtigt wird.

Für die weiteren Untersuchungen wird das Regressionsmodell mit dem höchsten Bestimmtheitsmaß  $b_{YX_1 \dots X_{k'}}$  verwendet. Als Maß für die Sensitivität eines Parameters  $X_j$  kann der Korrelationskoeffizient  $r_{YX_j}$  betrachtet werden, da er sich unmittelbar aus der Kovarianzmatrix, siehe (2.8), berechnen lässt. Taucht jedoch der Parameter  $X_j$   $l$ -mal in der multiplen Regressionsgleichung auf, weil z. B. quadratische und kubische Anteile von  $X_j$  zur Erklärung von  $Y$  notwendig sind, so berechnet sich  $r_{YX_j} = r_{YX_{j_1} \dots X_{j_l}}$  nach (2.9). Anhand der Größe der einzelnen Korrelationskoeffizienten  $r_{YX_j}$  lassen sich jetzt die Parameter sortieren und die wichtigsten selektieren. Im Balkendiagramm (siehe Abb. 3.5) ist das veranschaulicht. Um einen Überblick über die Zu-

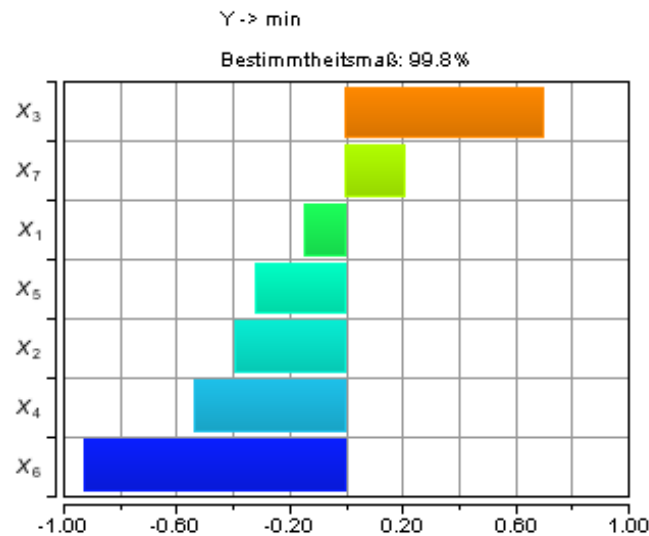


Abb. 3.5: Balkendiagramm mit Korrelationskoeffizienten zwischen  $Y$  und den  $X_i$ ,  $i = 1, \dots, 7$

sammenhänge im gesamten Optimierungsproblem zu gewinnen, empfiehlt sich die Darstellung der Pearson'schen Korrelationskoeffizienten in einem Korrelationsdiagramm (siehe Abb. 3.6). In den jeweiligen Feldern sind die Korrelationen zwischen allen Zielgrößen und Parametern, einschließlich den Korrelationen untereinander, dargestellt. Die farbliche Belegung der einzelnen Felder beschreibt analog wie im Balkendiagramm der Abb. 3.5 die Art und Stärke der Korrela-

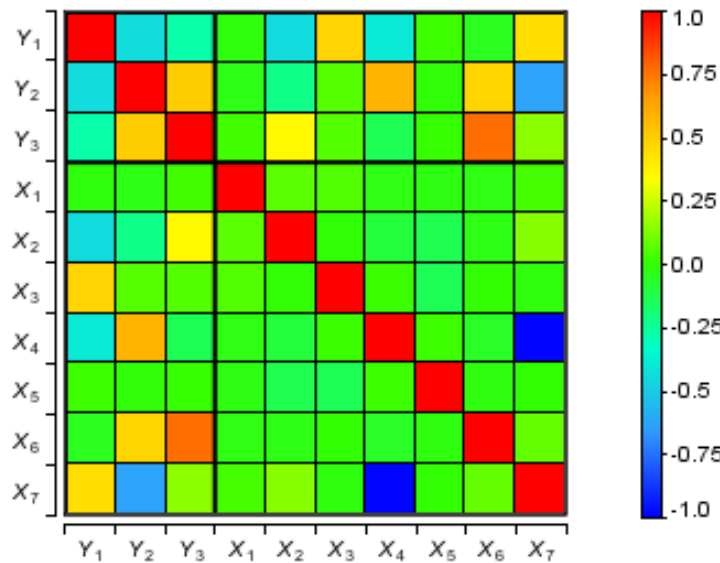


Abb. 3.6: Korrelationsdiagramm

tion. Auffallend sind hierbei die rot hervorgehobenen Korrelationen in der Hauptdiagonalen, da stets  $r_{Y_i Y_i} = r_{X_j X_j} = 1$  für alle  $i = 1, \dots, m$  und  $j = 1, \dots, k$  gilt. Aufgrund von Bemerkung 2.2 ist die Darstellung symmetrisch.

Fehlen jedoch signifikante Anteile des Parameters im Regressionsmodell, wird der Einfluss von ihm stark unterschätzt. Deshalb ist es bei der Korrelationsanalyse sehr wichtig, von einer gut bestimmten multiplen Regressionsgleichung auszugehen. Gilt für einen Parameter  $r_{YX_j} \approx 0$  und sind  $Y$  und  $X_j$  normalverteilt, so kann nach Bemerkung 2.2 auf die Unabhängigkeit zwischen beiden geschlossen werden. Bestätigt werden kann das durch den Kolmogoroff-Smirnov-Test, der die Normalverteilung von  $Y$  und  $X_j$  auf einen bestimmten Niveau sichert. Zudem muss geprüft werden, dass eventuelle Abweichungen zwischen  $r_{YX_j}$  und Null nicht signifikant sind. Dieses Vorgehen ist in [5] erläutert. Sind  $X_j$  und  $Y$  unabhängig, so gehört  $X_j$  nicht zu den sensitiven Parametern und braucht bei der Robustheitsbewertung (siehe Kapitel 4) bzgl.  $Y$  nicht berücksichtigt zu werden.

Eine Voraussetzung für die multiple Regression ist die Unabhängigkeit der Regressoren. Im Allgemeinen werden diese aber schwach korreliert sein. Das wird als Multikollinearität bezeichnet. Ein Grund dafür ist, dass die Stichprobe aus der Gesamtheit aller Lösungen stammt. Diese schwache Korrelation hat noch keinen gravierenden Einfluss auf die Genauigkeit der multiplen Regression. Kritisch ist es jedoch, wenn zwei Parameter im Regressionsmodell stark korreliert sind. Dann kann einer dieser beiden Parameter aus dem Regressionsmodell entfernt werden, wenn dabei das Bestimmtheitsmaß nicht wesentlich fällt.

Problematisch ist weiterhin die Selektion von Ausreißern, d.h. von Werten, die stark von allen anderen Werten hinsichtlich einer Einfluss- oder Zielgröße abweichen. Die manuelle Selektion von Ausreißern ist in nicht eindeutigen Fällen schwierig, da u.U. die Daten verfälscht werden. In [5] sind einige statistische Tests angegeben, nach denen unter bestimmten Voraussetzungen geprüft werden kann, ob es sich bei ausgewählten Werten um Ausreißer handelt. Ist ein Ausreißer ausfindig gemacht worden, so ist dieser noch vor der multiplen Regressionsanalyse aus der Stichprobe zu entfernen, da er erheblich die daraus resultierenden Korrelationskoeffizienten verfälschen kann. Aufgezeigt wird das in Abb. 3.7. Um diese Problematik zu umgehen, kann ein Korrelationskoeffizient (siehe 3.2.2) verwendet werden, der unempfindlicher gegenüber Ausreißern ist.

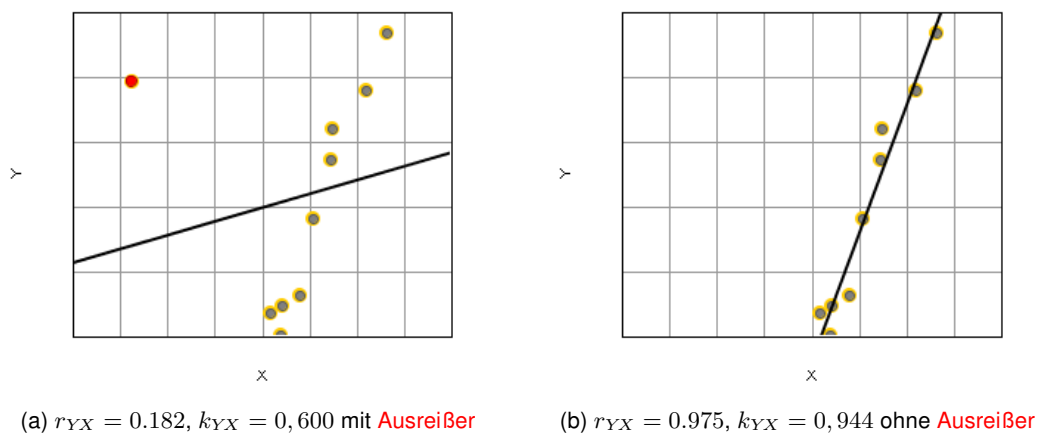


Abb. 3.7: Auswirkung von Ausreißern auf die Korrelationskoeffizienten  $r_{YX}$  und  $k_{YX}$

Ein weiteres Maß, um die Sensitivität eines Parameters zu charakterisieren, ist der Koeffizient relativer Wichtigkeit  $w_{YX_j}$ . Sei eine multiple Regression mit dem Bestimmtheitsmaß  $b_{YX_1 \dots X_k}$  gegeben. Dann ist

$$w_{YX_j} = b_{YX_1 \dots X_k} - b_{YX_1 \dots X_{j-1}X_{j+1} \dots X_k}.$$

Hierbei ist  $b_{YX_1 \dots X_{j-1}X_{j+1} \dots X_k}$  das Bestimmtheitsmaß der multiplen Regression, bei der der Parameter  $X_j$  entfernt wurde. Der Koeffizient relativer Wichtigkeit  $w_{YX_j}$  gibt an, um wie viel das Bestimmtheitsmaß  $b_{YX_1 \dots X_k}$  bei Weglassen des Parameters  $X_j$  fällt. Aus (3.2) folgt, dass  $0 \leq w_{YX_j} \leq 1$  ist. Werden die Parameter nach der Größe von  $w_{YX_j}$  geordnet, ergibt sich daraus eine Liste mit den einflussreichsten Parametern. Visualisiert wird das in Analogie zu Abb. 3.5 in einem Balkendiagramm. Jedoch werden aufgrund der Multikollinearität nicht immer alle sensitiven Parameter erkannt. Deshalb kann aus kleinen  $w_{YX_j}$  nicht gefolgert werden, dass sie keinen Einfluss auf  $Y$  haben. Im Allgemeinen lässt sich nur sagen, dass  $X_j$  unter gleichzeitiger Berücksichtigung der anderen Parameter keinen zusätzlichen signifikanten Beitrag zur Erklärung der Zielgröße  $Y$  leistet.

Einen Ausweg bietet hier der partielle Korrelationskoeffizient. Sei eine Zielgröße  $Y$  und die Parameter  $X_1$  bis  $X_k$  gegeben. Weiterhin sei

$$\mathbf{r}_{X_j U_j} = (r_{X_j X_i})_{\substack{i=1,\dots,k \\ i \neq j}}, \quad \mathbf{R}_{U_j U_j} = (r_{X_i X_l})_{\substack{l=1,\dots,k \\ l \neq j}}, \quad \text{und} \quad \mathbf{r}_{Y U_j} = (r_{Y X_i})_{i=1,\dots,k}$$

mit der Menge  $U_j = \{X_1, \dots, X_k\} \setminus \{X_j\}$  (siehe [4]). Der partielle Korrelationskoeffizient lautet

$$r_{Y X_j | U_j} = \frac{r_{Y X_j} - \mathbf{r}_{X_j U_j}^T \mathbf{R}_{U_j U_j}^{-1} \mathbf{r}_{Y U_j}}{\sqrt{\left(1 - \mathbf{r}_{X_j U_j}^T \mathbf{R}_{U_j U_j}^{-1} \mathbf{r}_{X_j U_j}\right) \left(1 - \mathbf{r}_{Y U_j}^T \mathbf{R}_{U_j U_j}^{-1} \mathbf{r}_{Y U_j}\right)}}.$$

Der partielle Korrelationskoeffizient  $r_{Y X_j | U_j}$  gibt den linearen Zusammenhang zwischen  $Y$  und  $X_j$  unter Partialisierung (Konstanthaltung) der Parameter  $X_i$ ,  $i \neq j$  an. Das bedeutet, dass die Einflüsse der  $X_i$  im Bezug auf die Korrelation zwischen  $Y$  und  $X_j$  ausgeschaltet werden. Dadurch können vorhandene Korrelationen  $r_{Y X_j}$  verschwinden, da  $X_j$  mit anderen, einflussreicheren Parametern korreliert ist, oder aber erst Hervortreten, weil  $r_{Y X_j}$  durch die anderen Parametereinflüsse verdeckt wird.

Hier wurden die jeweiligen Koeffizienten nur hinsichtlich einer Zielgröße betrachtet. Führt man die Analysen analog für die noch verbleibenden aus, können anschließend die wichtigsten Parameter der Optimierungsaufgabe herausgefiltert werden. Die Kenntnisse der sensitiven Parameter ist für ähnliche, darauf folgende Optimierungsprobleme von Vorteil, da beispielsweise eine a-priori Sensitivitätsanalyse entfallen kann.

### 3.2.2 Methode nach Kendall

Grundlage ist wieder eine gelöste MOA mit den Parametern  $X_1, \dots, X_k$  und den Zielgrößen  $Y_1, \dots, Y_m$ . Der Kendall'sche Rangkorrelationskoeffizient (vgl. [4]) ist prinzipiell nur für ordinal skalierte Merkmale, d.h. Merkmale, deren Ausprägungen eine Rangordnung besitzen, definiert. Eine Rangordnung könnte zwar ohne Weiteres eingeführt werden, z.B. durch die Sortierung der Zielfunktionswerte von  $Y_i$ , wobei der niedrigste dann den Rang eins erhält, jedoch geht dabei die Vergleichbarkeit mit den Korrelationen von  $X_j$  zu den anderen Zielgrößen verloren. Damit keine Rangordnung zwischen  $Y_i$  und  $X_j$  eingeführt werden muss, wird deshalb ein modifizierter Kendall'scher Korrelationskoeffizient  $k_{Y_i X_j}$  betrachtet. Ausgehend von  $Y_i$  und  $X_j$  wird das Produkt der Differenz zwischen der  $l$ - und  $m$ -ten paretooptimalen Lösung berechnet und

$$c_{lm}^{(ij)} = \begin{cases} 1, & \text{falls } (y_{il} - y_{im})(x_{jl} - x_{jm}) > 0 \\ 0, & \text{falls } (y_{il} - y_{im})(x_{jl} - x_{jm}) = 0 \\ -1, & \text{falls } (y_{il} - y_{im})(x_{jl} - x_{jm}) < 0 \end{cases}$$

gesetzt, für  $l = 1, \dots, n-1$  und  $m = l+1, \dots, n$ . Der Kendall'sche Korrelationskoeffizient  $k_{Y_i X_j}$  ist dann die normierte Summe der  $c_{lm}$ , d.h.

$$k_{Y_i X_j} = \frac{2}{n(n-1)} \sum_{l=1}^{n-1} \sum_{m=l+1}^n c_{lm}^{(ij)}. \quad (3.4)$$

Durch die Normierung liegt  $k_{Y_i X_j}$  ebenfalls im Bereich  $[-1, 1]$ . Ist  $k_{Y_i X_j}$  ca. 1 bzw. -1, so wächst bzw. fällt  $Y_i$  monoton bzgl.  $X_j$ . Aber die Art des Wachstums, ob beispielsweise linear oder exponentiell, wird durch  $k_{Y_i X_j}$  nicht vermittelt. In den meisten Fällen ist  $k_{Y_i X_j}$  kleiner als  $r_{Y_i X_j}$ , da durch die  $c_{lm}^{(ij)}$  nur das Vorzeichen, nicht aber der Abstand zwischen den Werten berücksichtigt wird. Dadurch ist der Kendall'sche Korrelationskoeffizient auch unempfindlicher gegenüber der Einwirkung von Ausreißern, was in Abb. 3.7 verdeutlicht wird. Sollte es vorkommen, dass  $r_{Y_i X_j}$  betragsmäßig kleiner ist als  $k_{Y_i X_j}$ , so kann auch ein Modellfehler bzgl.  $X_j$  die Ursache sein. Graphisch lässt sich  $k_{Y_i X_j}$  gut in Balkendiagrammen (siehe Abb. 3.5) darstellen.

### 3.2.3 Bewertung mit ANOVA

Ausgangspunkt ist hier wieder eine gut angepasste multiple Regression (vgl. 3.2.1) mit der Gleichung

$$\hat{y}(x_1, \dots, x_{k'}) = b_0 + b_1 x_1 + b_2 x_2 + \dots + b_{k'} x_{k'}$$

und der erklärten Varianz  $s_Y^2$ . Der  $j$ -te Parameter wird aus der Regressionsgleichung entfernt. Unter Umständen taucht dieser Parameter in  $q$  Gliedern auf. Für diese um den  $j$ -ten Parameter reduzierte neue Regressionsgleichung wird die erklärte Varianz  $s_{Yj}^2$  berechnet. Es ist leicht einzusehen, dass  $s_Y^2 \geq s_{Yj}^2$  gilt. Die Differenz zwischen beiden wird ins Verhältnis zur Restvarianz  $s^2$  der gegebenen Regression gesetzt. Durch die Skalierung mit dem Verhältnis aus  $n-1$  und  $q$  ergibt sich die Fisher'sche F-Statistik

$$F_j = \frac{n-1}{q} \cdot \frac{s_Y^2 - s_{Yj}^2}{s^2}.$$

Hohe Werte von  $F_j$  weisen auf einen großen Einfluss des Parameters  $X_j$  hin. Unter der Annahme, dass  $Y$  normalverteilt ist, kann geprüft werden, ob  $X_j$  signifikant notwendig zur Erklärung von  $Y$  ist. Gilt zum Niveau  $p$ , dass

$$F_j > c_{F(q, n-k'-1); 1-p}$$

ist, so muss an  $X_j$  zur Erklärung von  $Y$  festgehalten werden (vgl. [5]). Mit Hilfe von Satz 2.1 und Definition 2.16 wird  $F_j$  umgeformt.

$$\begin{aligned}
F_j &= \frac{n-1}{q} \cdot \frac{s_Y^2 - s_{Yj}^2}{s^2} \\
&= \frac{n-1}{q} \cdot \frac{b_{YX_1 \dots X_{k'}} s_Y^2 - b_{YX_{(1)} \dots X_{(k'-q)}} s_Y^2}{s^2} \\
&= \frac{n-1}{q} \cdot \frac{s_Y^2 (b_{YX_1 \dots X_{k'}} - b_{YX_{(1)} \dots X_{(k'-q)}})}{s^2} \\
&= \frac{n-1}{q} \cdot \frac{(n-k'-1) (b_{YX_1 \dots X_{k'}} - b_{YX_{(1)} \dots X_{(k'-q)}})}{(n-1) (1 - b_{YX_1 \dots X_{k'}})} \\
&= \frac{n-k'-1}{q} \cdot \frac{w_{YX_j}}{1 - b_{YX_1 \dots X_{k'}}}
\end{aligned}$$

Somit steht  $F_j$  im unmittelbaren Zusammenhang mit  $w_{YX_j}$ . Dadurch liefert  $F_j$  die gleichen Aussagen wie der Koeffizient relativer Wichtigkeit. Jedoch ist  $F_j \in (0, \infty)$  und damit für den Vergleich der Einflüsse von  $X_j$  auf unterschiedliche Zielgrößen ungeeignet. Bei der Sensitivitätsanalyse des Anwendungsbeispiels im Kapitel 5 wird deshalb die Bewertung mit der Varianzanalyse nicht mit betrachtet.

## 4 Robustheitsbewertung

Um sicher zu gehen, dass die aus der Optimierung einer MOA erhaltenen Lösungen den Anforderungen in der Praxis standhalten, werden diese Lösungen hinsichtlich ihrer Robustheit untersucht. Grund für die Unterschiede zu den berechneten Lösungen sind die in der Realität auftretenden Abweichungen in den Parametern. Deshalb ist die Bewertung der Optimierungsergebnisse hinsichtlich ihrer Robustheit von großer Bedeutung, damit die daraus resultierenden Eigenschaften der entstehenden Produkte gewährleistet werden können.

Die Robustheitsbewertung kann zu verschiedenen Zeitpunkten bei der Optimierung erfolgen. Vorteilhaft ist in allen Fällen, wenn die sensitiven Parameter bekannt sind, da die Parameter ohne Einfluss auf die Zielgrößen nicht mit berücksichtigt werden müssen. Im Abschnitt 4.1 wird die Bewertung von Lösungen nach bereits durchgeführter Optimierung der MOA vorgestellt. Darauf aufbauend kann für ausgewählte Lösungen eine Analyse der Robustheit mit den unter 4.2 vorgestellten Verfahren durchgeführt werden. Weiterhin ist es möglich, bereits bei der Optimierung die Robustheit der einzelnen Lösungen mit einfließen zu lassen, um als Ergebnis hauptsächlich robuste Optima zu erhalten. Ideen dazu werden im Abschnitt 4.3 betrachtet.

Unter Umständen kann eine robuste Lösung einer MOA ein oder mehrere schlechtere Zielfunktionswerte haben als andere optimale Lösungen. Dieser Kompromiss zwischen der Robustheit und den bestmöglichen Zielfunktionswerten ist häufig in Kauf zu nehmen. In Abb. 4.1 ist dieser Zusammenhang dargestellt.

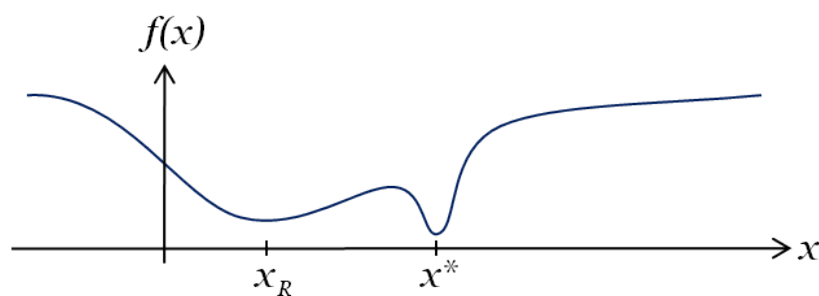


Abb. 4.1: Zu minimierende Zielfunktion mit globalen Optimum  $x^*$  und robusten Optimum  $x_R$

### 4.1 Direkte Auswertung

Um die Robustheit von Lösungen nach einer erfolgten Optimierung zu bewerten, werden im Folgenden zwei Möglichkeiten vorgestellt. Zum einen werden anhand von Streuungsmaßen, wie der

Varianz der Zielgrößen und der Parameter, Aussagen über die Robustheit einer Lösung generiert. Zum anderen wird mit prognostizierten Zielfunktionswerten aus einer multiplen Regression die Veränderung der Zielfunktionswerte einer Lösung analysiert und ausgewertet.

#### 4.1.1 Bewertung mittels Streuungsmaße

Betrachtet wird hier eine zulässige Lösung  $x$  einer MOA mit einer Umgebung  $U_x$ . Die Umgebung  $U_x$  setzt sich aus den einzelnen Schwankungsbereichen der Parameterwerte um  $x$  zusammen, d.h.

$$U_x = [x_1 - \varepsilon_1, x_1 + \varepsilon_1] \times [x_2 - \varepsilon_2, x_2 + \varepsilon_2] \times \dots \times [x_n - \varepsilon_n, x_n + \varepsilon_n]$$

mit  $\varepsilon_i > 0, \forall i = 1, \dots, n$ . Alle in  $U_x$  befindlichen zulässigen Lösungen werden nun zur Bewertung der Robustheit von  $x$  herangezogen. Die Anzahl der Lösungen in  $U_x$  sei  $k$ . Unter Umständen kann  $k$  sehr klein sein, so dass die folgenden statistischen Analysen an Aussagekraft verlieren oder nicht durchführbar sind. Empfehlenswert ist es hierbei, die Untersuchung nur für die sensiblen Parameter der MOA durchzuführen, da dadurch  $k$  größer werden kann. Im Folgenden sei  $k \gg 2$  und die Lösungen in der Umgebung  $U_x$  gleichmäßig verteilt. Für jede Zielgröße wird nun mit statistischen Größen, wie der Varianz, der Bereich untersucht, in dem die Zielfunktionswerte der  $k$  Lösungen aus  $U_x$  streuen. Dabei werden die Mittelwerte, Varianzen, Spannweiten und empirischen Verteilungen der Zielfunktionswerte der einzelnen Zielgrößen betrachtet. Die Robustheit von  $x$  wird danach bewertet, ob die Varianzen der Zielgrößen klein genug sind und wie die Zielfunktionswerte innerhalb der Spannweite verteilt sind. Eine geringe Varianz durch starke Einschränkung der Schwankungsbereiche der Parameter zu erzielen, ist statistisch nicht sinnvoll. In Abb. 4.2 sind verschiedene Verteilungen dargestellt. Grün hervorgehoben ist dabei der

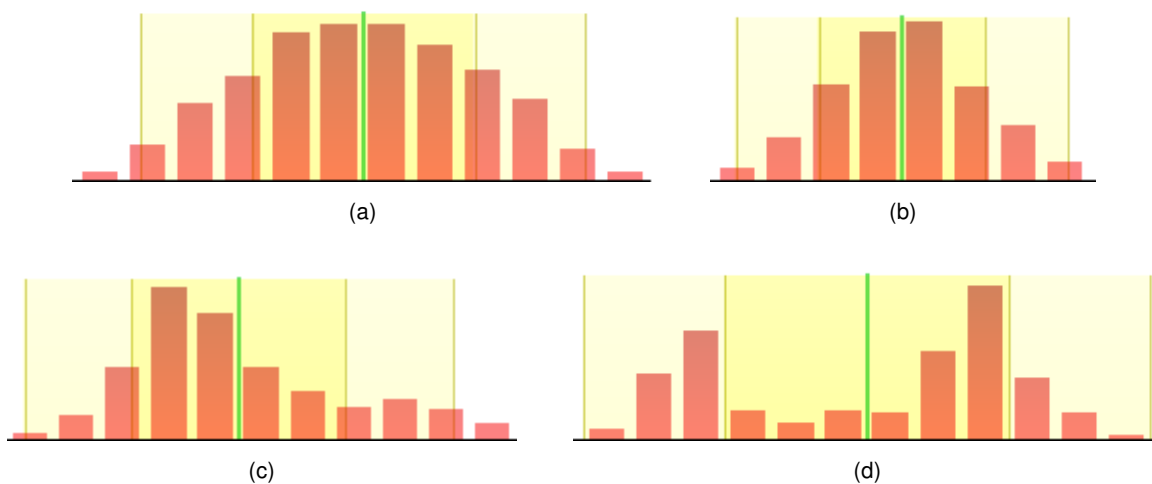


Abb. 4.2: Histogramme zur empirischen Verteilung der Zielgrößen



Zielfunktionswert  $f_i(\mathbf{x})$ . Bei den Darstellungen 4.2a und 4.2b sind die Zielfunktionswerte symmetrisch um  $f_i(\mathbf{x})$  verteilt, wobei  $f_i(\mathbf{x})$  in der Nähe des arithmetischen Mittelwertes liegt. Die Varianz bei Abb. 4.2b ist wesentlich kleiner als bei 4.2a. Schiefe Verteilungen, wie in Abb. 4.2c, sind Anzeichen für schleichende Verschlechterungen der Zielgrößen. Treten mehrgipflige Verteilungen der Zielfunktionswerte auf, siehe Abb. 4.2d, ist die Robustheit der Lösung  $\mathbf{x}$  als schlecht zu bewerten. Aus diesen Erkenntnissen gilt es, Maße zur Bewertung der Robustheit einer Lösung  $\mathbf{x}$  zu generieren. Sei  $s^2$  die Varianz einer Zielgröße  $Y$  bezogen auf die Umgebung  $U_{\mathbf{x}}$  und  $y = f(\mathbf{x})$  der Zielfunktionswert der zu untersuchenden Lösung. Aufgrund der Schwerpunkteigenschaft des arithmetischen Mittels gilt

$$s^2 = \frac{1}{k-1} \sum_{i=1}^k (y_i - \bar{y})^2 \leq \frac{1}{k-1} \sum_{i=1}^k (y_i - y)^2 = s_Y^2.$$

Daraus folgt, dass

$$\frac{s_Y^2}{s^2} \geq 1$$

ist. Unter Annahme einer eingipfligen und annähernd symmetrischer Verteilung der Zielfunktionswerte aus  $U_{\mathbf{x}}$  ist mit  $\frac{s_Y^2}{s^2} - 1$  nahe 0 die Lösung  $\mathbf{x}$  robust. Dabei bleibt die Variation der Parameterwerte in  $U_{\mathbf{x}}$  vorerst unberücksichtigt. Um das zu kompensieren, wird gleichzeitig das Verhältnis der Variationskoeffizienten

$$\psi_{YX_j} = \left| \frac{v_Y}{v_{X_j}} \right|$$

zwischen der Zielgröße und den Parametern bezogen auf  $U_{\mathbf{x}}$  betrachtet. Auffällig für nicht robuste Lösungen sind kleinere Variationskoeffizienten der Parameter als die der Zielgrößen, d.h. wenn

$$\psi_{YX_j} > 1$$

gilt. Voraussetzung für diese Annahme ist, dass die Parameter und Zielgrößen ausschließlich positive oder negative Werte, aber nicht beide zugleich annehmen. Dadurch ist auch gesichert, dass die jeweiligen Mittelwerte verschieden 0 sind. Um zwei Lösungen hinsichtlich ihrer Robustheit zu vergleichen, wird aus eigener Überlegung heraus der Robustheitsindex

$$\iota_{\mathbf{x}} = \sum_{i=1}^m \left( \left( \frac{s_{Y_i}^2}{s_i^2} - 1 \right) \sum_{j=1}^n \psi_{Y_i X_j} \right)$$

eingeführt. Unter den Bedingungen, dass die Varianzen der Zielgrößen  $Y_i$  klein genug, die Zielfunktionswerte entsprechend Abb. 4.2b um  $f_i(\mathbf{x})$  verteilt,  $k$  hinreichend groß ist und die Lösungen in der Umgebung gleichmäßig verteilt sind, liefert der Robustheitsindex  $\iota_{\mathbf{x}}$  eine Aussage über das Varianz- und Variationsverhalten der Lösung  $\mathbf{x}$ . Je kleiner  $\iota_{\mathbf{x}}$  ist, umso robuster wird die Lösung  $\mathbf{x}$  angesehen. Eine Lösung  $\mathbf{x}$  kann als robust bewertet werden, falls  $\iota_{\mathbf{x}} \leq m \cdot n$  ist. Diese obere Schranke entsteht durch die Abschätzung von  $\psi_{Y_i X_j}$  und  $\frac{s_{Y_i}^2}{s_i^2} - 1$  mit jeweils 1. Für in den Parameterwerten ähnliche Lösungen sind auch die Robustheitsindizes ähnlich, da sich die jeweiligen

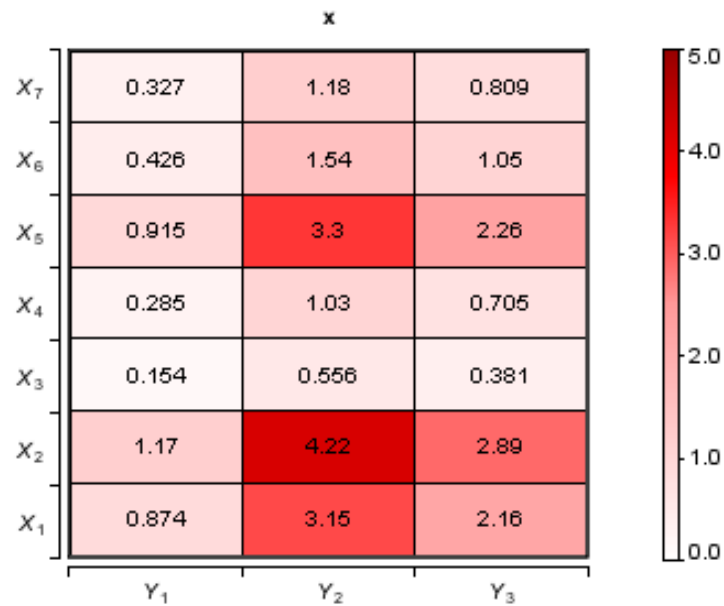


Abb. 4.3: Beispielhafte Darstellung der Variationsverhältnisse  $\psi_{Y_i X_j}$  einer Lösung  $x$  einer MOA

Umgebungen überschneiden. Sind bei der Untersuchung einer Lösung die genannten Bedingungen verletzt, kann anhand des Robustheitsindex keine Aussage über ihre Robustheit getroffen werden. Die einzelnen Variationsverhältnisse  $\psi_{Y_i X_j}$  können dennoch wichtige Informationen liefern, wenn für einzelne Parameter und Zielgrößen die vorhin erwähnten Bedingungen gelten. Veranschaulicht werden die Variationsverhältnisse beispielsweise in Diagrammen wie bei Abb. 4.3. Darin sind die Parameter gut zu erkennen, die die Robustheit von  $x$  am meisten gefährden (in diesem Fall  $X_2$ ).

#### 4.1.2 Bewertung durch Prognose der Zielfunktionswertänderungen

Grundlage für diese Untersuchung ist eine gegebene, gut angepasste multiple Regressionsgleichung  $\hat{y}_i$  für jede einzelne Zielgröße. Diese wird analog wie bei der Sensitivitätsanalyse bestimmt. Untersucht wird hierbei die Differenz zwischen dem tatsächlichen Zielfunktionswert der zu untersuchenden Lösung  $x$  und dem geschätzten Zielfunktionswert der um einen Vektor  $\Delta x$  veränderten Lösung  $x$ , also

$$f_i(x) - \hat{y}_i(x + \Delta x) .$$

Die Wahl der Komponenten von  $\Delta x$  kann hierbei beliebig erfolgen. Im Weiteren wird aber die  $j$ -te Komponente  $\Delta x_j$  prozentual zur Länge des jeweiligen Parameterintervalles  $l_{X_j}$  gewählt, d.h.  $\Delta x_j = p \cdot l_{X_j}$  mit  $p \in [-1, 1]$ . Dadurch können die Auswirkungen der Änderungen von Parameterwerten untereinander besser verglichen werden. Die Zielfunktionswertänderungen werden dabei einzeln für jede Parameteränderung und gesamt bei Berücksichtigung aller Parameter-

änderungen betrachtet. Graphisch dargestellt werden die einzelnen Daten in Diagrammen wie bei Abb. 4.4. Die farbliche Hinterlegung der einzelnen „Kästchen“ gibt dabei Auskunft über die

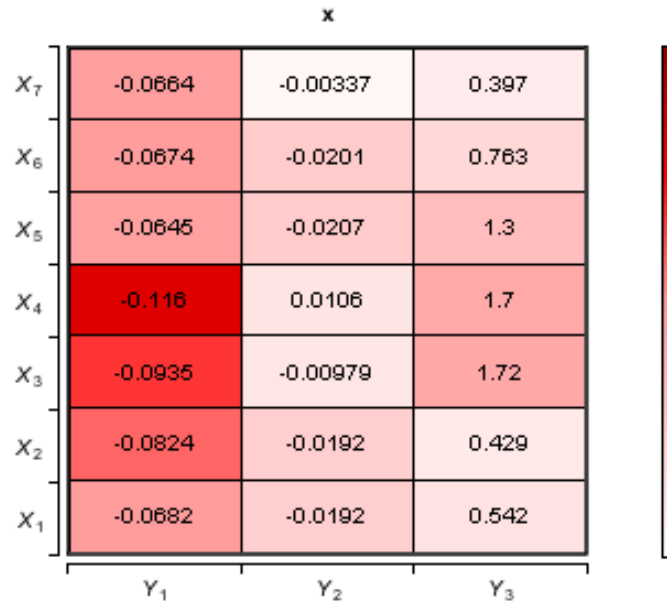


Abb. 4.4: Graphische Darstellung der Zielfunktionswertänderungen bei Variation einer Lösung  $x$

Art und Stärke der Veränderung des  $i$ -ten Zielfunktionswertes bei Variation des  $j$ -ten Parameters. Außerdem ist die Farbe an die Größenordnung der Zielgröße angepasst. Die Spaltensumme entspricht der Zielfunktionswertänderung bei Berücksichtigung aller Werteänderungen in den Parametern ausgehend von der Lösung  $x$ . Die Robustheit von  $x$  wird nach dem Einvernehmen mit den geschätzten Zielfunktionswertänderungen bewertet. Hieraus können auch für einflussreiche Parameter Toleranzintervalle abgeschätzt werden, um die Zielfunktionswerte auf einem bestimmten Niveau zu halten. Um mehrere Lösungen hinsichtlich ihrer Robustheit direkt vergleichen zu können, wird im Folgenden ein Maß zur Bewertung einer Lösung  $x$  konstruiert. Sei  $\Delta y_{ij}$  die Änderung des  $i$ -ten Zielfunktionswertes bei Abweichung des  $j$ -ten Parameters um  $\Delta x_j$ . Analog wie bei der Berechnung von  $\Delta x_j$  sei  $\Delta y_i = p \cdot l_{y_i}$ , d.h.  $\Delta y_i$  ist ein prozentualer Anteil an der Länge des Wertebereiches der  $i$ -ten Zielfunktion. Der Betrag des Verhältnisses

$$\left| \frac{\Delta y_{ij}}{\Delta y_i} \right|$$

gibt an, wie die prognostizierten Zielfunktionswertänderungen in Relation zu einer anteiligen Abweichung  $\Delta y_i$  des jeweiligen Zielfunktionswertes stehen. Der Durchschnitt

$$\bar{d}_i = \frac{1}{k} \sum_{j=1}^k \left| \frac{\Delta y_{ij}}{\Delta y_i} \right|$$

aus diesen Verhältnissen, wobei  $k$  die Anzahl der Parameter ist, gibt die mittlere relative Abweichung des  $i$ -ten Zielfunktionswertes bezogen auf den Bereich  $[y_i - \Delta y_i, y_i + \Delta y_i]$  an. Damit wird

der Robustheitsgrad

$$\eta_i = e^{-\bar{d}_i}$$

einer Lösung  $x$  definiert. Ist die mittlere relative Abweichung  $\bar{d}_i$  nahe 0, wird die Robustheit der Lösung  $x$  mit ca. 1 bewertet. Wächst hingegen  $\bar{d}_i$ , so sinkt auch die Robustheit. Bei  $\bar{d}_i = 1$ , d.h. wenn die einzelnen Zielfunktionswertänderungen rund den prozentualen Anteil an der Länge des Intervalles der  $i$ -ten Zielfunktion entsprechen, wird die Robustheit der Lösung  $x$  mit  $\eta_i = 0,37$  bewertet. Für  $\bar{d}_i \rightarrow \infty$  ist der Robustheitsgrad 0. Eine Lösung  $x$  heißt bezogen auf eine Zielgröße  $i$  robust, wenn  $\eta_i \geq 0,7$  gilt. Um die Vergleichbarkeit mit anderen Lösungen zu ermöglichen, kann entweder ein Mittelwert oder das Minimum beziehungsweise Maximum aus allen  $\eta_i$  gebildet werden. In Abb. 4.5 sind die  $\eta_i$  einer Lösung  $x$  aus einer MOA graphisch dargestellt.

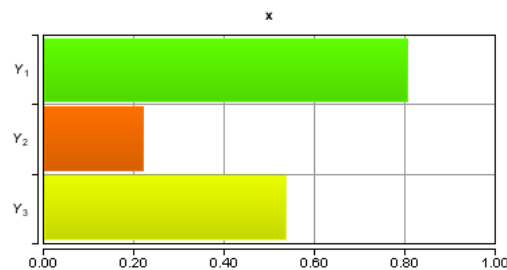


Abb. 4.5: Balkendiagramm der Robustheitsgrade  $\eta_i$  einer Lösung  $x$  aus einer MOA

Problematisch hierbei ist, wie bereits bei der Sensitivitätsanalyse, die Konstruktion einer gut angepassten multiplen Regression für jede Zielgröße. Die Prognose der Zielfunktionswertänderungen ist nur für Parametervariationen innerhalb der gegebenen Parametergrenzen sinnvoll.

## 4.2 Aufbauende Verfahren

Bestehen noch Unsicherheiten in der Robustheitsbewertung einer Lösung  $x$ , da beispielsweise die Bedingungen für die Bewertung in den vorangegangenen Abschnitten nicht erfüllt sind, werden mit Hilfe von mathematischen Modellen der Stochastik neue Lösungen generiert und die dazugehörigen Zielfunktionswerte berechnet. Die mit Abweichungen behafteten Parameter werden hierbei mit normalverteilten Zufallsgrößen modelliert. Der Mittelwert  $\mu_j$  der  $j$ -ten Zufallsgröße beziehungsweise des Parameters wird auf  $x_j$  gesetzt. Die Varianzen der normalverteilten Parameter sind entsprechend ihren geschätzten Abweichungen festzulegen. Weiterhin sind vor Erzeugung der Lösungen Intervalle für die Zielgrößen zu bestimmen, z.B. mit den Mittelpunkten  $f_i(x)$  plus/minus einer akzeptablen Toleranz. Die Stichprobe aus den Parameterraum wird hier mit einer Monte-Carlo Simulation erzeugt. Möglich wäre auch die Generierung der Lösungen mit dem Latin-Hypercube Sampling (vgl. [8]), jedoch kann es dabei ungewollt zu Korrelationen

zwischen den Parametern kommen. Ausgehend davon, dass  $k$  Lösungen mit der Monte-Carlo Simulation erzeugt werden sollen, werden in Abhängigkeit der den Parametern zugewiesenen Normalverteilungen  $k$  Lösungen mit einem Zufallsgenerator ermittelt. Anschließend werden die Zielfunktionswerte der  $k$  generierten Lösungen berechnet. In Abb. 4.6 ist das Ergebnis einer solchen Simulation für den zweidimensionalen Fall dargestellt. Rot hervorgehoben ist der Punkt der

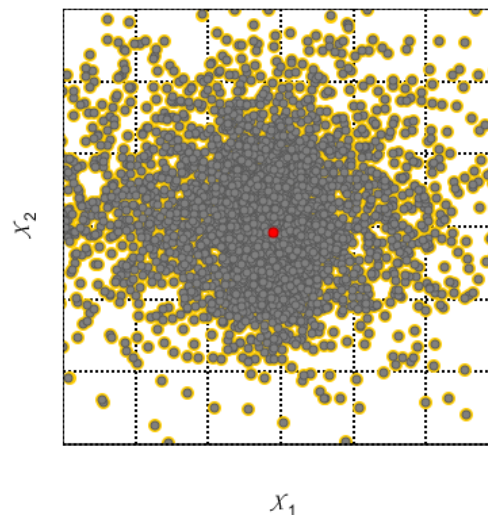


Abb. 4.6: Monte-Carlo Simulation zweier normalverteilter Zufallsgrößen

jeweiligen Mittelwerte. Die Anzahl der darin dargestellten Lösungen ( $k = 4000$ ) ist für praxisrelevante Fälle viel zu groß. Gute Ergebnisse werden bereits erzielt, wenn  $k = n^2$  ist, wobei  $n$  die Anzahl der betrachteten Parameter ist. Anhand dieser generierten Stichprobe wird die Robustheitsbewertung der Lösung  $x$  mit den unter Abschnitt 4.1 vorgestellten Methoden untersucht. Der wesentliche Unterschied zu diesen Methoden ist, dass durch die vor Beginn der Simulation festgelegten Intervalle der Zielgrößen die Versagenswahrscheinlichkeit, d.h. dass  $x$  außerhalb dieser liegt, berechnet werden kann. Das kann beispielsweise durch Ermittlung der relativen Häufigkeit geschehen. Diese und weitere Ideen zur Bewertung der Robustheit von  $x$  mit Wahrscheinlichkeiten entstammen der Zuverlässigkeitstheorie und sind in [8] näher erläutert. Hier wird darauf nicht näher eingegangen, da der entstehende Aufwand durch die erneute Berechnung von Zielfunktionswerten nur in seltenen Fällen gerechtfertigt ist.

### 4.3 Robuste Optimierung

Erste Ansätze zur robusten Optimierung entstanden bereits Mitte des 20. Jahrhunderts. Das Ziel dabei ist, ein robustes Optimum, d.h. eine zulässige Lösung der MOA, die unempfindlich gegenüber kleinen Variationen in den Parameterwerten ist, zu finden (siehe Abb. 4.1).

Neben verschiedenen Methoden aus der Theorie der Strukturzuverlässigkeit stammt einer der bekanntesten Ansätze zur robusten Optimierung von Taguchi (vgl. [17]), der sich mit Produkt- und Prozessentwicklung befasst. Ziel dabei ist es, die Varianz der Zielgrößen zu minimieren. Das soll nicht durch kleine Schwankungsbereiche der Parameterwerte erreicht werden. Es gilt vielmehr, eine Unempfindlichkeit der Zielfunktionswerte gegenüber den streuenden Eingangsgrößen zu erzielen. Zu beachten ist auch, dass das Konvergenzverhalten angewendeter Optimierungsverfahren nicht eingeschränkt wird. Als Kenngröße für die Robustheit des Optimierungsprozesses verwendet Taguchi die Signal-Rausch-Verhältnisse

$$S/N_i = 10 \log \left( \frac{\bar{y}_i^2}{s_{Y_i}^2} \right),$$

wobei  $\bar{y}_i$  der Mittelwert und  $s_{Y_i}^2$  die Varianz der  $i$ -ten Zielgröße ist. Soll nur die Varianz der Zielgrößen während des Verlaufes der Optimierung betrachtet werden, so kann das Robustheitsmaß

$$RM_i = 10 \log \left( \frac{1}{s_{Y_i}^2} \right)$$

für jede Zielgröße  $i$  verwendet werden. Diese Maße können nun an geeigneter Stelle in ein beliebiges Optimierungsverfahren integriert werden.

Eine weitere Möglichkeit die robuste Optimierung im Lösungsverfahren zu integrieren, ist die vorhandenen Zielfunktionen  $f_i(\mathbf{x})$  in neue Zielfunktionen

$$F_i(\mathbf{x}) = \alpha_i f_i(\mathbf{x}) + \beta_i f_{R_i}(\mathbf{x})$$

umzuwandeln, die aus einen Optimierungs- und Robustheitsteil bestehen. Der Robustheitsteil  $f_{R_i}(\mathbf{x})$  ist eine über die Reihenentwicklung nach Taylor bestimmte Funktion, die die Krümmung der Zielfunktion im Punkt  $\mathbf{x}$  beschreibt. Die Wichtung von  $f_i(\mathbf{x})$  und  $f_{R_i}(\mathbf{x})$  erfolgt mit den Faktoren  $\alpha$  und  $\beta$ . In Abhängigkeit vom verwendeten Optimierungsverfahren sind unter Umständen die Optimalitätsbedingungen anzupassen, damit die Konvergenz des Verfahrens gesichert bleibt.

## 4.4 Bewertungsablauf

Eine Robustheitsuntersuchung wird je nach verwendeter Methode unterschiedlich durchgeführt. Die Bewertungsmethode mit den Streuungsmaßen ist von den zwei vorgestellten direkten Verfahren die aufwendigere, da bei ihr die Voraussetzungen für jede einzelne Lösung überprüft werden müssen. Empfehlenswert ist dabei die Vorgehensweise, dass zuerst die Lösung mit den Robustheitsindex bewertet wird und dann erst die Voraussetzungen überprüft werden, falls sich der dadurch entstehende Aufwand rechtfertigt. Bei der Methode mit den Zielfunktionswertänderungen empfiehlt es sich, zuerst die einzelnen Robustheitsgrade der Lösung zu analysieren und

dann mittels der graphischen Darstellung der einzelnen Zielfunktionswertveränderungen die Robustheit der Lösung detailliert zu betrachten. Zu gewährleisten ist hierbei, dass die grundlegende multiple Regression gut bestimmt ist. Bei den aufbauenden Verfahren ist der Ablauf der Bewertung nach der Generierung einer neuen Lösungsmenge, je nach verwendeter Bewertungsmethode, wie oben beschrieben durchzuführen. Die Ermittlung von Ausfallwahrscheinlichkeiten erfolgt dabei parallel zur Berechnung der anderen Robustheitsmaße. Bei Verfahren der robusten Optimierung kann eine nachfolgende Robustheitsbewertung eingespart werden, wenn die daraus erhaltenen Ergebnisse hinsichtlich der Robustheit zufriedenstellend sind.





## 5 Technisches Anwendungsbeispiel

Die EngineeringToolbox der IAV GmbH beinhaltet das Modul UNO, das zur Lösung von technischen Optimierungsproblemen verwendet wird. Im Modul UNO steht eine Optimierungsbibliothek bereit, die mit verschiedenen Tools aus der ET verknüpft werden kann. Dadurch kann die IAV GmbH alle anwendungsspezifischen Problemstellungen optimieren, deren Simulationen mit den Tools in der ET realisiert sind.

### 5.1 Problemstellung

Die folgende Problemstellung stammt aus der Anwendung V-CD (Virtual Chain Drive) der EngineeringToolbox und wird in [10] beschrieben. Sie dient unter anderem zur kinetostatischen und geometrischen Auslegung von Kettentrieben bei der Motorsteuerung und wurde in der IAV GmbH entwickelt. Die CAD-Oberfläche der Anwendung V-CD ist in Abb. 5.1 zusehen. Darin befindet

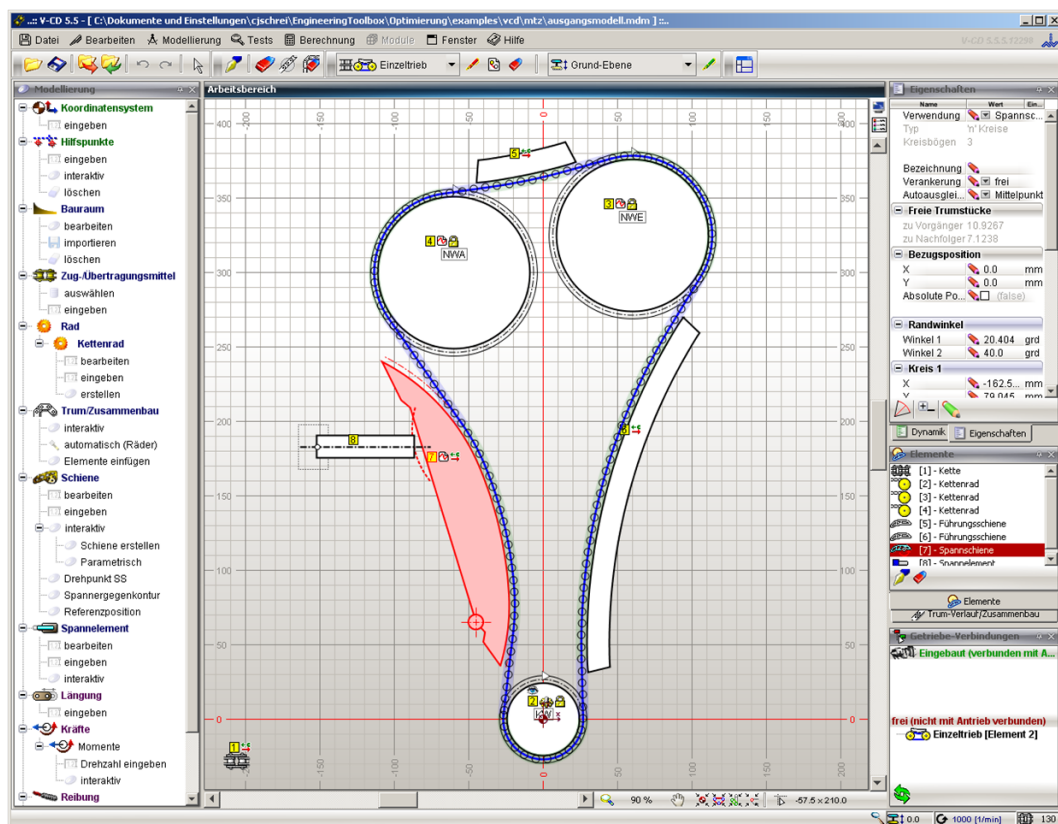


Abb. 5.1: Simulation von Steuerketten durch V-CD

sich auch die Darstellung des zu optimierenden Kettentriebs. Der Einsatz von Ketten als Steuertrieb wird bei der Entwicklung von modernen Motoren immer häufiger. In diesem Beispiel dient der Kettentrieb zur Steuerung zweier Nockenwellen eines Motors. Angetrieben wird der Steuertrieb über ein auf der Kurbelwelle befindliches Zahnrad. Um zu gewährleisten, dass die Kette des Steuertriebes, die auch Trum genannt wird, weder überspringt noch den Kontakt zu den zwei Führungsschienen und der Spannschiene (rot hervorgehoben in Abb. 5.1) verliert, muss sie vorgespannt werden. Die entstehende Kraft in der Trum heißt Vorspannkraft. Das Spannsystem dient nicht nur zur Spannung der Kette, sondern dämpft ihre Schwingungen und reguliert die Länge des Trums, die sich mit der Zeit einstellt.

Das Optimierungsproblem beinhaltet zwölf Parameter und sechs Zielgrößen. In Abb. 5.2 sind die geometrischen Beziehungen zwischen den Einflussgrößen dargestellt und in Tab. 5.1 erläutert.

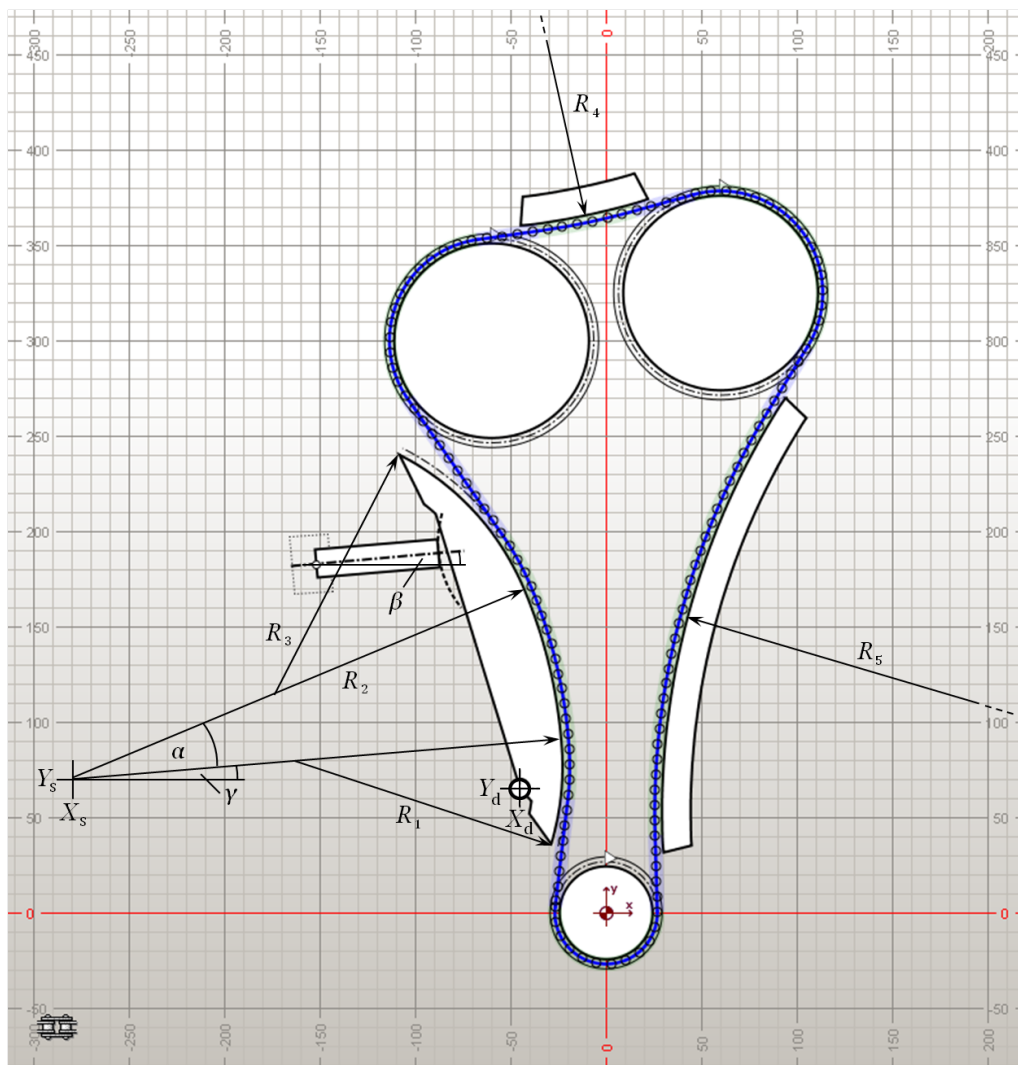


Abb. 5.2: Geometrische Darstellung der Parameter des Steuertriebes

Parameter	Bereich	Beschreibung
$X_d$	$[-50, -45]$	x-Koordinate des Drehpunktes der Spannschiene
$Y_d$	$[45, 70]$	y-Koordinate des Drehpunktes der Spannschiene
$X_s$	$[-330, -280]$	x-Koordinate des Scheitelpunktes vom Hauptradius der Spannschiene
$Y_s$	$[50, 90]$	y-Koordinate des Scheitelpunktes vom Hauptradius der Spannschiene
$R_1$	$[140, 190]$	Einlaufradius der Spannschiene
$R_2$	$[260, 310]$	Hauptradius der Spannschiene
$R_3$	$[135, 180]$	Auslaufradius der Spannschiene
$R_4$	$[250, 450]$	Hauptradius der oberen Führungsschiene
$R_5$	$[400, 600]$	Hauptradius der rechten Führungsschiene
$\alpha$	$[10, 40]$	Öffnungswinkel der Spannschiene
$\beta$	$[5, 15]$	Bezugsrichtung des Spannelementes
$\gamma$	$[-17, 17]$	Bezugsrichtung der Spannschiene

Tab. 5.1: Beschreibung der Parameter des Anwendungsbeispiels

Die Werte der Parameter sind bis auf die drei Winkelmaße alle in Millimetern angegeben.

Zielgröße	Beschreibung
H	Kolbenhub des Spannelementes zwischen vorgespannter und maximaler Auslenkung
V	absolute Abweichung zwischen dem Verhältnis von Vorspannkraft zu Spannerkraft und 2, 5
$V_v$	absolute Abweichung zwischen den Verhältnissen von Vorspannkraft zu Spannerkraft im gelängten und ungelängten Zustand
$P_r$	Reibleistung in der Trum und an den Kontaktflächen
T	Summe aller quadratischen Abweichungen zwischen den freien Trumlängen und vier Kettengliedern
U	Kontaktlänge über alle Ein- und Auslaufkreisbögen

Tab. 5.2: Beschreibung der Ziele des Anwendungsbeispiels

Bei der Optimierung des Kettentriebs werden die in Tab. 5.2 beschriebenen Ziele verfolgt. Jede einzelne Zielgröße gilt es hierbei zu minimieren. Die Zielgrößen H, V,  $V_v$  und  $P_r$  dienen zur Optimierung der technischen Anforderungen. Zur gleichzeitigen Gewährleistung einer funktionalen Geometrie sind die Ziele T und U in die Optimierungsaufgabe integriert.

## 5.2 Auswertung

Die Lösung des Optimierungsproblems erfolgte mit dem Modul UNO in der EngineeringToolbox. Als Optimierungsverfahren wurde ein probabilistischer Genetischer Algorithmus verwendet, der auf dem Genetischen Algorithmus MrBOA (Multiobjective real-coded Bayesian Optimization Algorithm, siehe [1]) basiert. Dieser Algorithmus wurde verwendet, da die Paretomenge gut approximiert wird. Als Ergebnis hat das Lösungsverfahren eine Paretomenge mit 3755 Elementen geliefert.

### 5.2.1 Sensitivitätsanalyse

Für jede Zielgröße wurde eine multiple Regressionsgleichung ermittelt. Jede dieser multiplen Regressionsgleichungen ist linear in allen Parametern, d.h. der Graph der jeweiligen Regression ist eine Hyperebene. In Tab. 5.3 sind die Bestimmtheitsmaße dargestellt. Da diese ausreichend

	H	V	$V_v$	$P_r$	T	U
Bestimmtheitsmaß in %	95,0	92,0	89,9	83,2	96,1	76,9

Tab. 5.3: Bestimmtheitsmaße der multiplen Regressionen

hoch sind, können nun gute Aussagen bezüglich der Sensitivitätsanalyse getroffen werden. Zuerst wird das Korrelationsdiagramm aus Abb. 5.3 betrachtet. Dadurch erhält man einen ersten Eindruck über die Zusammenhänge in der Lösungsmenge des Optimierungsproblems. Auffällig sind hierbei die Korrelationskoeffizienten nahe Null zwischen allen Zielgrößen und den Parametern  $R_1$ ,  $R_4$  und  $X_d$ . Daraus kann geschlossen werden, dass diese drei Parameter nicht zu den sensitiven Einflussgrößen in der gesamten Optimierungsaufgabe zählen. Die starken negativen Korrelationen  $r_{HV}$  und  $r_{VP_r}$  deuten auf die Ziele hin, die am meisten untereinander konkurrieren. Zudem ist aus dem Korrelationsdiagramm ersichtlich, dass neben  $R_2$  und  $X_s$  auch  $\alpha$  und  $\gamma$  stark negativ korreliert sind. Aus diesen starken Korrelationen lässt sich folgern, dass der Spielraum für die optimale Position der Spannschiene sehr schmal ist. Die anderen Korrelationskoeffizienten (nach Bravais-Pearson) zwischen den Parametern sind überwiegend im grünen Bereich, d.h. die Parameter sind höchstens schwach korreliert.

Die sensitiven Parameter werden nun für die Zielgröße H ermittelt. Dabei werden die vorgestellten Sensitivitätsmaße, außer die Bewertung nach der Varianzanalyse, miteinander verglichen. In Abb. 5.4 sind die vier unterschiedlichen Koeffizienten in Balkendiagrammen dargestellt. Beim Vergleich des Pearson'schen mit dem Kendall'schen Korrelationskoeffizienten fällt hierbei auf, dass der Kendall'sche Korrelationskoeffizient kleiner ist als der Pearson'sche, aber den gleichen

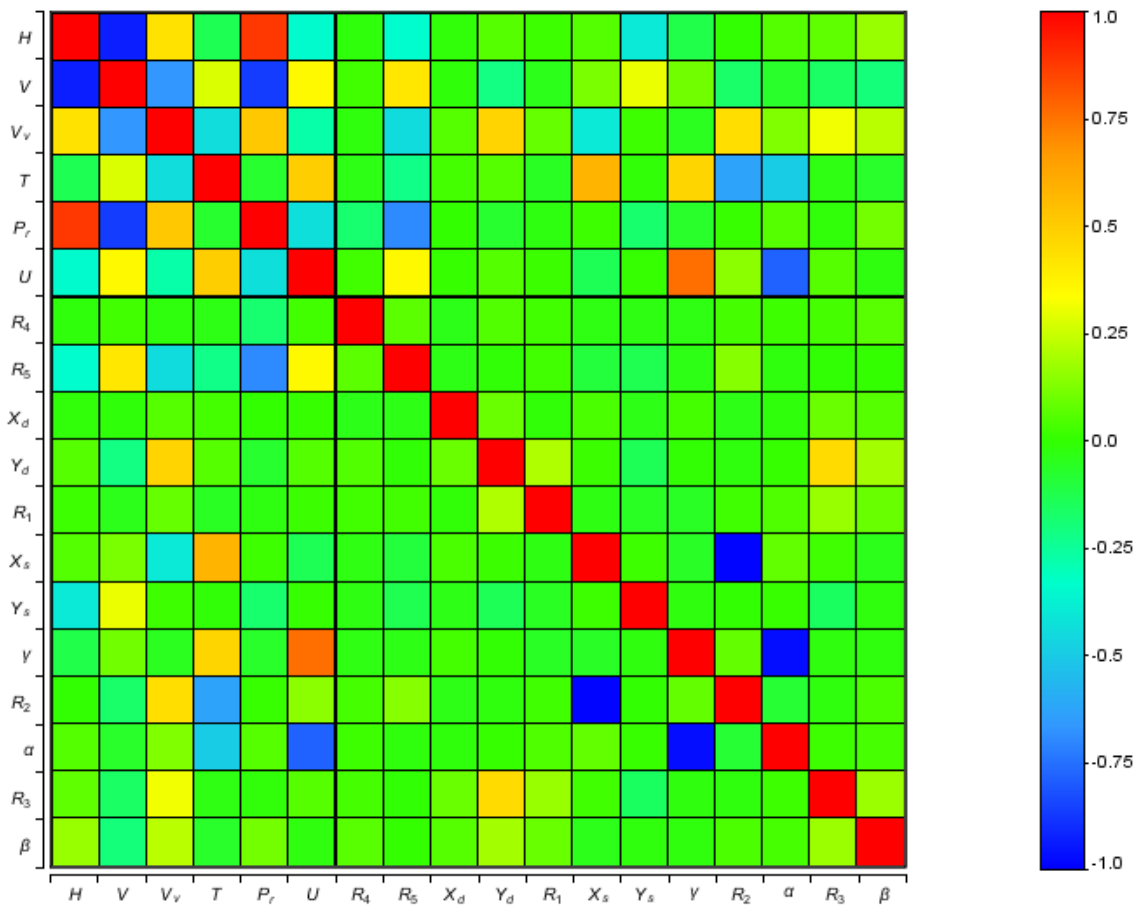


Abb. 5.3: Korrelationsdiagramm zum Anwendungsbeispiel

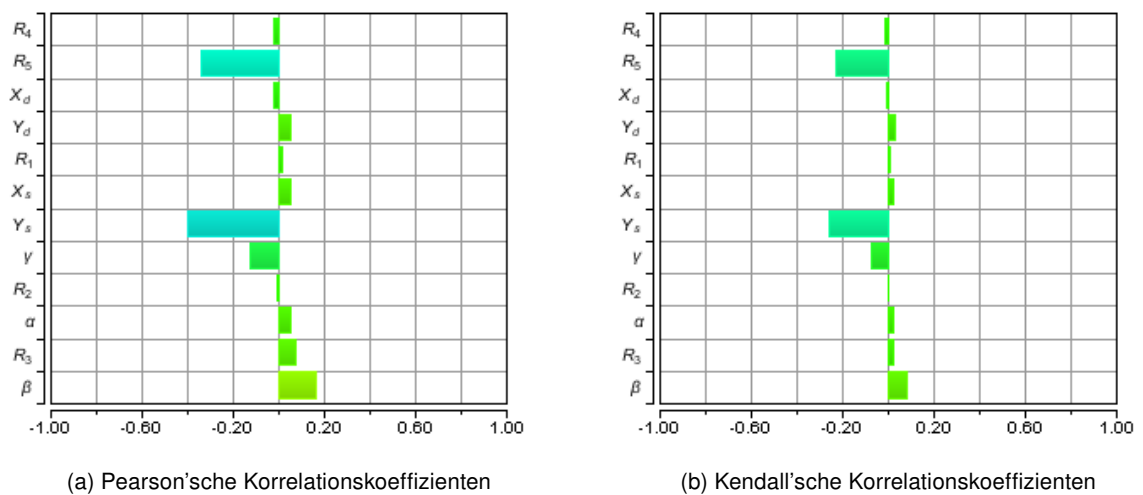


Abb. 5.4: Korrelationskoeffizienten zwischen der Zielgröße H und allen Parametern

Trend besitzt. Als sensitive Parameter für die Zielgröße H kommen dabei  $R_5$  und  $Y_s$  in Betracht. Jedoch sind  $R_5$  und  $Y_s$  nur schwach mit H korreliert. Betrachtet man zusätzlich den partiellen

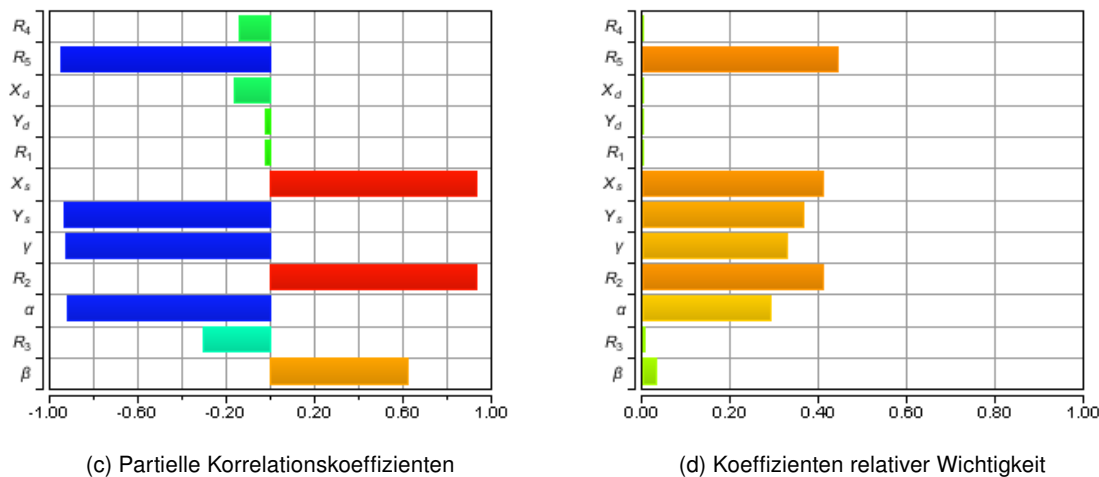


Abb. 5.4: Korrelationskoeffizienten zwischen der Zielgröße H und allen Parametern

Korrelationskoeffizienten zwischen allen Parametern und H, fällt auf, dass nicht nur  $R_5$  und  $Y_s$  Einfluss auf H haben. Erklären lässt es sich damit, dass diese Einflüsse durch alle Korrelationen der Parameter mit H verdeckt werden. Weiterhin ist zu beobachten, dass der Koeffizient relativer Wichtigkeit die gleichen Parameter wie der partielle Korrelationskoeffizient als sensitiv kennzeichnet. Das muss aber nicht immer der Fall sein. Im Ermessen des Anwenders liegt auch, ob der Parameter  $\beta$  mit zu den einflussreichsten Parametern gezählt wird oder nicht. Analog zu dieser Vorgehensweise werden die sensitiven Parameter für die anderen Zielgrößen bestimmt. Die Ergebnisse zur Sensitivitätsanalyse des Anwendungsbeispiels sind in Tab. 5.4 zusammengefasst.

Parameter	Zielgrößen					
	H	V	$V_v$	T	$P_r$	U
$R_4$					?	
$R_5$	✓	✓	✓		✓	✓
$X_d$						
$Y_d$		?	✓			
$R_1$						
$X_s$	✓	✓	✓	✓	✓	
$Y_s$	✓	✓			✓	
$\gamma$	✓	✓		✓	✓	✓
$R_2$	✓	✓	✓	✓	✓	
$\alpha$	✓	✓		✓	✓	✓
$R_3$						
$\beta$	?	?			?	

Tab. 5.4: Zusammenfassung der Sensitivitätsanalyse des Anwendungsbeispiels

Dabei sind die sensitiven Parameter bezüglich einer Zielgröße mit einem Häkchen (✓) gekennzeichnet. Die Parameter, die zusätzlich sensitiv sein können, sind durch ein Fragezeichen (?) hervorgehoben. Alle anderen Parameter haben folglich keinen erkennbaren, wesentlichen Einfluss auf die Zielgrößen. Gut zu erkennen ist in der Tab. 5.4, dass die am meisten konkurrierenden Ziele, d.h.  $H$ ,  $V$  und  $P_r$ , von der gleichen Parametermenge abhängen. Die insgesamt 20 weiteren Balkendiagramme mit den Korrelationskoeffizienten zu den anderen Zielgrößen befinden sich unter A.2.

### 5.2.2 Robuste Lösungen

Aus der Paretomenge wurden drei unterschiedliche Lösungen herausgefiltert, für die die Robustheitsbewertung durchgeführt wird. Die Zielfunktionswerte sind in Tab. 5.5 dargestellt. Der

Lösung	Zielfunktionswerte					
	$H$	$V$	$V_v$	$T$	$P_r$	$U$
$x_1$	9,17	0,32	0,72	51,01	46,73	128,26
$x_2$	9,38	0,24	0,78	31,48	46,80	121,13
$x_3$	9,63	0,20	0,76	72,29	48,63	126,54

Tab. 5.5: Zielfunktionswerte der ausgewählten Lösungen

Entscheidungsgraph in Abb. 5.5 liefert eine gute Übersicht über die Verteilung der Zielfunktionswerte der Lösungen. Zuerst wird die Robustheit der drei Lösungen mittels den Streuungsmaßen

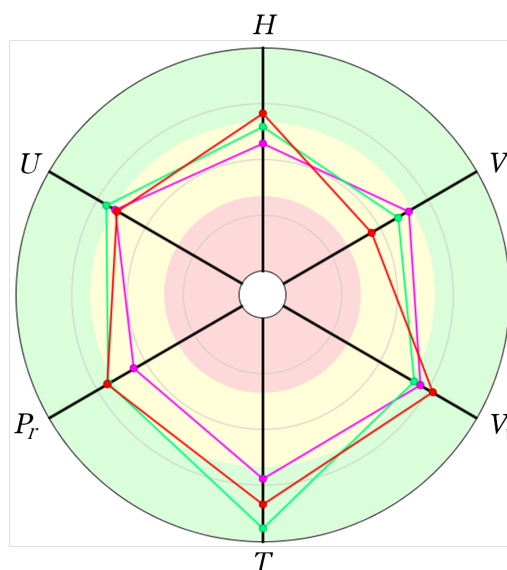


Abb. 5.5: Entscheidungsgraph mit den Lösungen  $x_1$ ,  $x_2$  und  $x_3$

bewertet. Dabei werden nur die sensitiven Parameter der Optimierungsaufgabe betrachtet (vgl. Tab. 5.4). Das heißt, die Parameter  $R_4$ ,  $X_d$ ,  $R_1$ ,  $R_3$  und  $\beta$  fließen nicht mit in die Robustheitsbewertung der Lösungen ein. Die Schwankungsbereiche, aus denen die jeweiligen Umgebungen hervorgehen, wurden derart gewählt, dass  $\epsilon_j$  10% der Länge der Parameterintervalle um den  $j$ -ten Parameter entspricht. In Tab. 5.6 ist die Anzahl  $k$  der Lösungen in der Umgebung, die

Lösung	$k$	$s_1$	$s_2$	$s_3$	$s_4$	$s_5$	$s_6$	$\iota$
$x_1$	14	0,46	0,11	0,06	32,96	1,45	12,83	36,32
$x_2$	2	-	-	-	-	-	-	-
$x_3$	22	0,62	0,15	0,07	18,06	1,93	10,15	4,90

Tab. 5.6: Robustheitsbewertung mittels Streuungsmaße

Standardabweichungen  $s_i$  der  $i$  Zielfunktionswerte in Bezug auf die jeweilige Umgebung und der Robustheitsindex  $\iota$  abgebildet. Dabei sind die Zielfunktionswerte der Lösungen  $x_1$  und  $x_3$  annähernd symmetrisch um die Ausgangswerte verteilt. Für die Lösung  $x_2$  kann keine Aussage zur Robustheit getroffen werden, da  $k$  zu klein ist. Die anderen Lösungen  $x_1$  und  $x_3$  sind robust, da  $\iota$  jeweils kleiner gleich 42 (6 Zielgrößen multipliziert mit 7 Parameter) ist. Da  $k$  aber relativ klein ist, besitzt der Robustheitsindex für beide keine hohe Aussagekraft. In der Abb. 5.6 sind

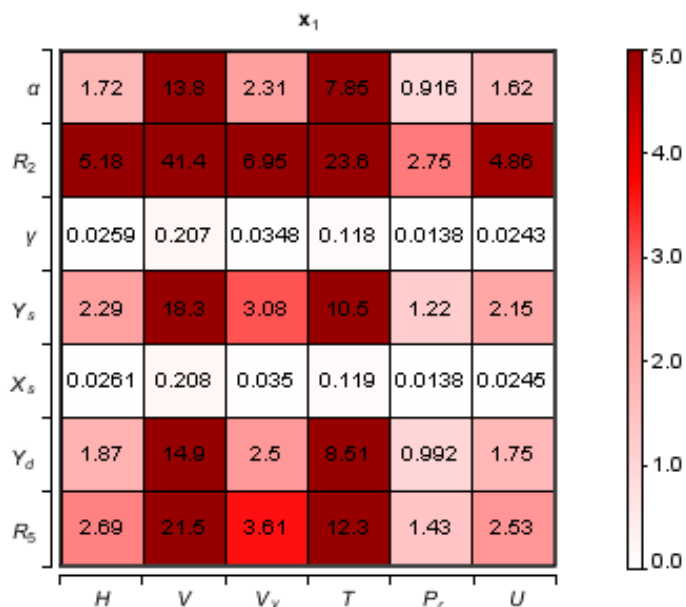
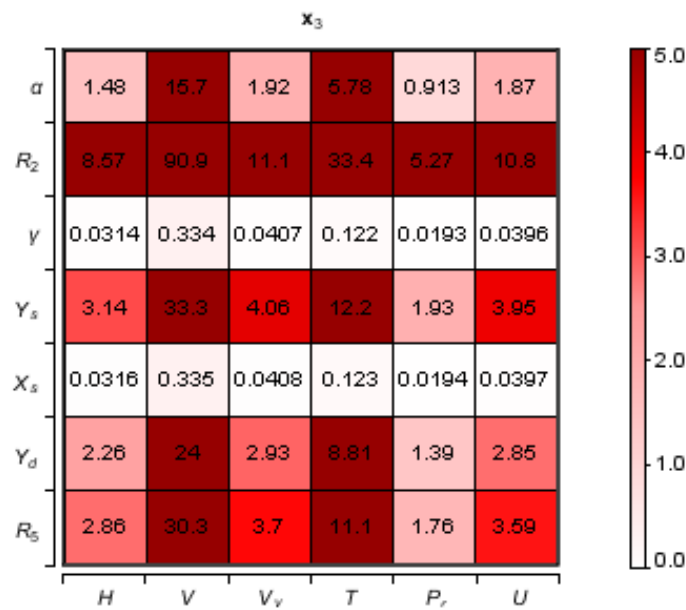


Abb. 5.6: Variationsverhältnisse  $\psi$  der Lösungen  $x_1$  und  $x_3$

die Variationsverhältnisse  $\psi$  für die Lösungen  $x_1$  und  $x_3$  dargestellt. Beim direkten Vergleich der beiden Abbildungen fällt auf, dass fast alle Variationsverhältnisse der eigentlich „robusteren“ Lö-



Abb. 5.6: Variationsverhältnisse  $\psi$  der Lösungen  $x_1$  und  $x_3$ 

sung  $x_3$  größer sind als die von  $x_1$ . Zum einen liegt das daran, dass  $k$  verhältnismäßig klein ist. Zum anderen liegen die Zielfunktionswerte von  $x_3$  zentraler in den Zielfunktionswerten aus der Umgebung als die von  $x_1$ . Auffallend bei beiden Lösungen ist, dass trotz der vielen roten Bereiche in den Darstellungen der Variationsverhältnisse, die Lösungen als robust bewertet werden. Ursache dafür sind die nicht gleichmäßig verteilten Lösungen in den jeweiligen Umgebungen.

Die Bewertung der drei gegebenen Lösungen erfolgt jetzt durch die Prognose der Zielfunktionswertveränderungen. Zugrunde liegen hierbei die gleichen, gut bestimmten multiplen Regressionen mit allen Parametern wie bei der Sensitivitätsanalyse (siehe auch Tab. 5.3). Weiterhin ist  $p = 0, 1$ . Die dazugehörigen Diagramme der Zielfunktionswertveränderungen sind in Abb. 5.7 zu sehen. Die einzelnen Robustheitsgrade  $\eta$  sind in Tab. 5.7 aufgelistet. Dabei ist die Gesamtbe-

Lösung	$\eta_H$	$\eta_V$	$\eta_{V_v}$	$\eta_T$	$\eta_{P_r}$	$\eta_U$	Gesamtbewertung
$x_1$	0,296	0,264	0,437	0,649	0,427	0,773	0,474
$x_2$	0,369	0,110	0,681	0,673	0,491	0,370	0,449
$x_3$	0,253	0,265	0,375	0,626	0,353	0,812	0,447

Tab. 5.7: Robustheitsbewertung mittels Zielfunktionswertänderungen

wertung das arithmetische Mittel der einzelnen Robustheitsgrade jeder Lösung. Hierbei werden alle drei Lösungen durch den Robustheitsgrad annähernd identisch bewertet. Die Lösung  $x_1$  ist in Bezug zur Gesamtbewertung die Beste aus den drei Ausgewählten. Bei der vorhergehenden

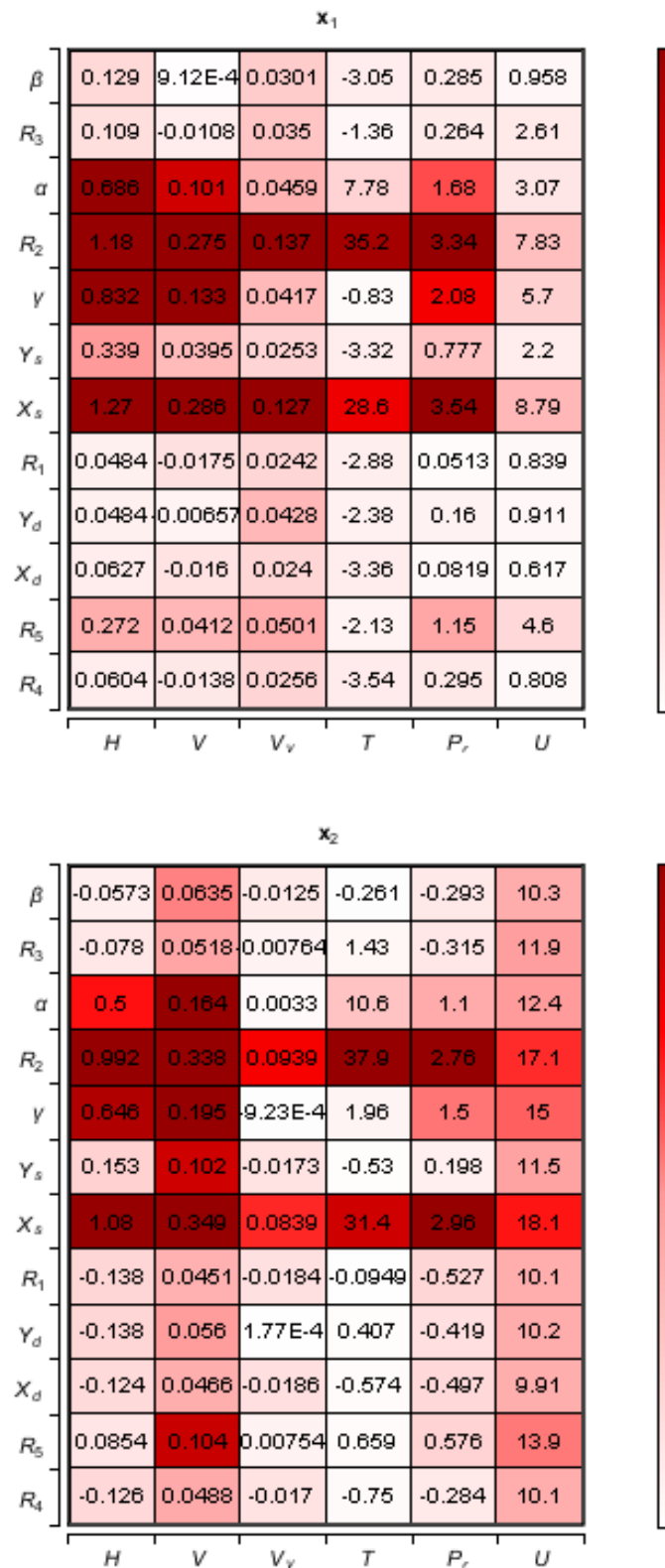


Abb. 5.7: Zielfunktionswertveränderungen der drei Lösungen

Methode war das nicht der Fall. In den Diagrammen der Zielfunktionswertveränderungen ist gut zu erkennen, dass bei allen drei Lösungen die Parameter  $X_s$ ,  $R_2$ ,  $\alpha$  und  $\gamma$  die Robustheit am meis-

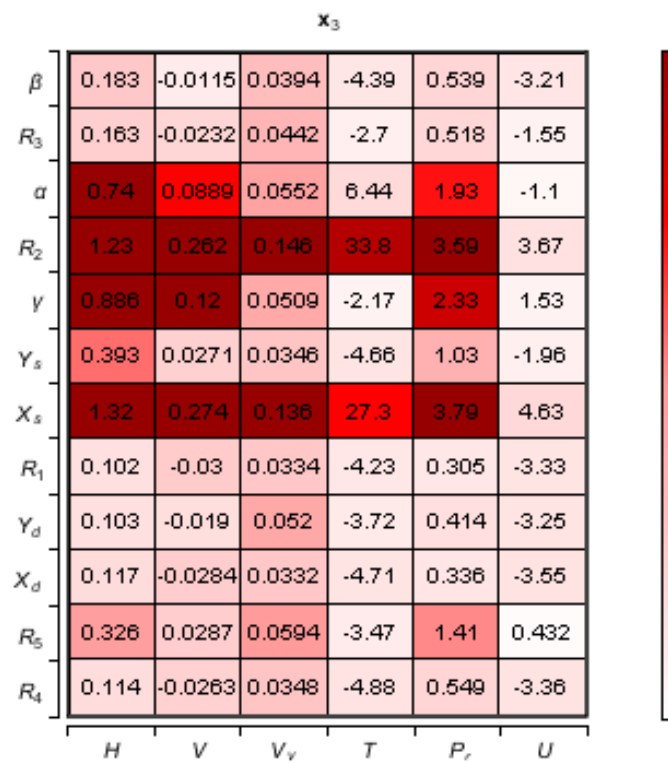


Abb. 5.7: Zielfunktionswertveränderungen der drei Lösungen

ten gefährden. Dadurch gilt für alle drei Lösungen, dass die Toleranzen der oben aufgezählten Parameter entsprechend festzulegen sind, falls die Zielfunktionswertveränderungen nicht akzeptabel sind. Das gilt jedoch nur, wenn eine der Lösungen verwendet wird. Zudem decken sich diese Ergebnisse mit der Sensitivitätsanalyse, da diese Parameter zu den einflussreichsten der Optimierungsaufgabe gehören.



## 6 Schluss

### 6.1 Zusammenfassung

Die vorliegende Bachelorarbeit befasst sich mit Methoden zur Sensitivitätsanalyse und Robustheitsbewertung für multikriterielle Optimierungsverfahren. Nach der Erläuterung wichtiger Grundlagen wird die Sensitivitätsanalyse zu unterschiedlichen Zeitpunkten im Optimierungsprozess betrachtet. Neben der a-priori Analyse, in der durch numerische Verfahren Aussagen über die einflussreichsten Parameter der MOA zu Beginn des Lösungsprozesses generiert werden, sind bei der a-posteriori Sensitivität verschiedene Methoden betrachtet worden, die sensitiven Parameter anhand der erhaltenen Lösungsmenge der MOA mittels statistischen Kenngrößen zu erkennen. Jede vorgestellte Methode hat dabei Stärken und Schwächen. Der Korrelationskoeffizient nach Bravais-Pearson liefert die Einflüsse der einzelnen Parameter in den meisten Fällen sehr exakt. Voraussetzung dafür ist jeweils eine gut angepasste multiple Regression der Zielgrößen durch die Parameter. Schwierig dabei ist das Erfassen von eventuellen Ausreißern, da diese die Ergebnisse des Korrelationskoeffizienten nach Bravais-Pearson stark verfälschen können. Der Korrelationskoeffizient nach Kendall ist dagegen unempfindlicher gegenüber den Verfälschungen durch Ausreißer, liefert aber nur eine Aussage zur Stärke des Trends des Einflusses der Parameter auf die Zielgrößen. Unter Umständen sind die Einflüsse einiger Parameter durch starke Korrelationen anderer Parameter zur jeweiligen Zielgröße verdeckt. Dabei hilft der partielle Korrelationskoeffizient, da er den Einfluss anderer Parameter ausschaltet und dadurch weitere Parametereinflüsse herausgefiltert werden können. Jedoch erkennt der partielle Korrelationskoeffizient nur lineare Einflüsse der Parameter. Durch den Koeffizienten relativer Wichtigkeit werden in den Regel die wichtigsten sensitiven Parameter aufgedeckt. Dabei kann es aber vorkommen, dass weniger einflussreiche Parameter nicht identifiziert werden. Die Methode nach ANOVA liefert die gleichen Ergebnisse wie der Koeffizient relativer Wichtigkeit. Abschließend lässt sich sagen, dass je nach Problemstellung die eine oder andere Analysemethode bessere Ergebnisse erzielt. Generell liefert der Korrelationskoeffizient nach Bravais-Pearson für die meisten technischen Probleme die besten Ergebnisse.

Die Robustheitsbewertung von Lösungen aus multikriteriellen Optimierungsaufgaben ist wichtig für die Gewährleistung der Eigenschaften der entstehenden Produkte. Die Kenntnisse über die sensitiven Parameter des Optimierungsproblems sind hierbei von Vorteil, da die Robustheitsbewertung auf diese Parameter beschränkt werden kann. Die hier vorgestellten Verfahren zur Bewertung der Robustheit von Lösungen liefern unter den jeweiligen Voraussetzungen aussagekräftige Ergebnisse. Die Methode mit den Streuungsmaßen ist die aufwendigere der zwei vor-

gestellten Verfahren zur direkten Bewertung der Lösungen, da viele unterschiedliche statistische Kennwerte betrachtet werden müssen. Die Bewertung der Lösungen mit der Prognose der Zielfunktionswertänderungen basiert, ähnlich wie bei der Sensitivitätsanalyse, auf gut angepassten, multiplen Regressionen. Da die zwei direkten Methoden unterschiedliche Voraussetzungen haben, ist fast immer eine der beiden Bewertungsmethoden anwendbar. Andernfalls können Verfahren verwendet werden, die eine neue Lösungsmenge in Bezug zur untersuchenden Lösung konstruieren und daran die Robustheitsbewertung der Lösung durchführen. Weiterhin werden Ideen der robusten Optimierung vorgestellt. Durch ein modifiziertes Optimierungsverfahren oder Hinzufügen verschiedener Robustheitsmaße zur Optimierungsaufgabe wird versucht, robuste Lösungen zu konstruieren.

Geeignete Verfahren zur Sensitivitäts- und Robustheitsbewertung für multikriterielle Optimierungsverfahren wurden schließlich im Modul MCA der IAV-EngineeringToolbox implementiert und auf das in dieser Arbeit vorgestellte Beispiel angewendet. Die jeweils beschriebenen graphischen Darstellungen zur besseren Auswertung der Ergebnisse wurden ebenfalls im Modul MCA integriert. Dadurch steht den Ingenieuren der IAV-GmbH ein umfangreiches Angebot zur Auswertung multikriterieller Optimierungsergebnisse bezüglich Sensitivität und Robustheit zur Verfügung.

## 6.2 Ausblick

Eine wesentliche Grundlage für einige Methoden zur Sensitivitäts- und Robustheitsbewertung bilden gut bestimmte multiple Regressionen. Um die Güte der Regressionen auch bei nichtlinearen Problemen zu gewährleisten, sollten echt nichtlineare multiple Regressionen betrachtet werden. Zudem ist es möglich, den kompletten Modellfindungsprozess für die multiplen Regressionen zu automatisieren. Trotz der Komplexität der Automatisierung, ist das ein lohnenswerte Methode zur Entwicklung intelligenter Auswertungsverfahren. Weiterhin sind andere Verfahren zur Sensitivitätsanalyse und Robustheitsbewertung denkbar. Zur Sensitivitätsanalyse sind Methoden und Verfahren des Data Minings geeignet, da diese Eigenschaften und Muster in großen Datenmengen erkennen. Die Integration der Sensitivitätsanalyse in Optimierungsverfahren, d.h. diese Verfahren entscheiden während des Optimierungsprozesses selbstständig anhand bestimmter Sensitivitätsmaße, welche Parameter relevant für die Optimierung sind, ist sehr empfehlenswert, da z.B. die Rechenzeit gesenkt werden kann. Kombiniert mit Ideen der robusten Optimierung entstehen somit komplexe Optimierungsverfahren, die den Auswertungsprozess der erhaltenen Ergebnisse erheblich verkürzen. Die Umsetzung anderer Methoden zur Robustheitsbewertung, wie z.B. die Ermittlung von Versagenswahrscheinlichkeiten, ist ebenfalls empfehlenswert. Dadurch kann die Robustheit von Lösungen vielfältig gesichert werden.

# A Anhang

## A.1 Verwendete Formeln

### A.1.1 Wichtige Verteilungen

Eine oft benötigte Verteilung ist die  $\chi^2$ -Verteilung. Seien  $\mathcal{U}_i$  unabhängige, identisch  $N(0, 1)$  verteilte ZG. Dann ist

$$\mathcal{Y} = \sum_{i=1}^n \mathcal{U}_i^2$$

$\chi^2$ -verteilt mit  $n$  Freiheitsgraden, d.h.  $\mathcal{Y} \sim \chi^2(n)$ . Der Erwartungswert einer  $\chi^2$ -verteilten ZG  $\mathcal{Y}$  ist  $n$  und die Varianz beträgt  $2n$ . Für  $\mathcal{Y} \sim \chi^2(n)$  gilt außerdem

$$\frac{\mathcal{Y} - n}{\sqrt{2n}} \xrightarrow{n \rightarrow \infty} N(0, 1) .$$

Die Verteilung der ZG

$$\mathcal{T} = \frac{\mathcal{U}}{\sqrt{\frac{\mathcal{Y}}{n}}} \sim t(n),$$

entspricht der Student'schen  $t$ -Verteilung, wenn  $\mathcal{U} \sim N(0, 1)$  und  $\mathcal{Y} \sim \chi^2(n)$  ist. Die ZG  $\mathcal{T}$  ist dann  $t$ -verteilt mit  $n$  Freiheitsgraden. Der Erwartungswert  $E\mathcal{T} = 0$  für  $n > 1$  und  $D^2\mathcal{T} = \frac{n}{n-2}$  für  $n > 2$ . Eine Eigenschaft einer  $t$ -verteilten ZG  $\mathcal{T}$  ist:

$$\mathcal{T} \xrightarrow{n \rightarrow \infty} N(0, 1) .$$

Als letztes wird die Fisher'sche  $F$ -Verteilung betrachtet. Seien  $\mathcal{Y}_n \sim \chi^2(n)$  und  $\mathcal{Y}_m \sim \chi^2(m)$  gegeben. Die ZG

$$\mathcal{F} = \frac{\frac{\mathcal{Y}_n}{n}}{\frac{\mathcal{Y}_m}{m}} \sim F(n, m)$$

ist  $F$ -verteilt mit  $n$  und  $m$  Freiheitsgraden. Der Erwartungswert  $E\mathcal{F} = \frac{n}{n-2}$  für  $n > 2$  und die Varianz  $D^2\mathcal{F} = \frac{2n^2(n+m-2)}{m(n-2)^2(n-4)}$  bei  $n > 4$ .

### A.1.2 Statistische Lageparameter und Streuungsmaße

Weitere Lageparameter einer Stichprobe sind:

1. empirischer Median:

$$\bar{x}_{0.5} = \begin{cases} x_{\frac{n+1}{2}}, & \text{falls } n \text{ ungerade} \\ \frac{1}{2}(x_{\frac{n}{2}} + x_{\frac{n}{2}+1}), & \text{falls } n \text{ gerade} \end{cases}$$

2. geometrische Mittel:

$$\bar{x}_g = \sqrt[n]{\prod_{i=1}^n x_i}$$

3. empirischer Modalwert:

$$\bar{x}_M = x_i \text{ mit der größten Häufigkeit}$$

Andere Streuungsmaße der Lageparameter einer Stichprobe bzgl. eines Merkmals X sind:

1. empirische Standardabweichung:

$$s_X = +\sqrt{s_X^2}$$

2. empirischer Variationskoeffizient:

$$v_X = \frac{s_X}{\bar{x}}, \bar{x} \neq 0$$

3. mittlere absolute Abweichung vom Median:

$$m_X = \frac{1}{n} \sum_{i=1}^n |x_i - \bar{x}_{0.5}|$$

4. Spannweite:

$$l_X = \max_i x_i - \min_i x_i$$

Weitere statistische Kennwerte sind (analog zu den Definitionen 2.10 und 2.11):

1. Schiefe:

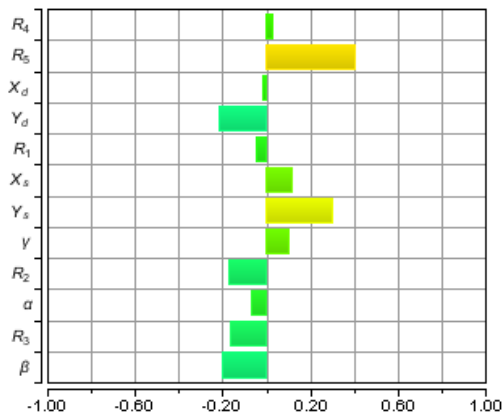
$$g_1 = \frac{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^3}{\sqrt{\left(\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2\right)^3}}$$



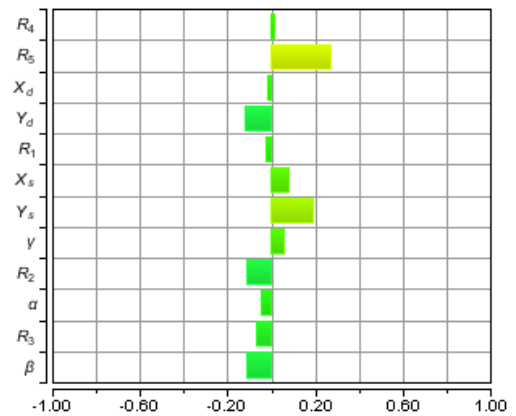
2. Exzess:

$$g_2 = \frac{n \sum_{i=1}^n (x_i - \bar{x})^4}{\left( \sum_{i=1}^n (x_i - \bar{x})^2 \right)^2} - 3$$

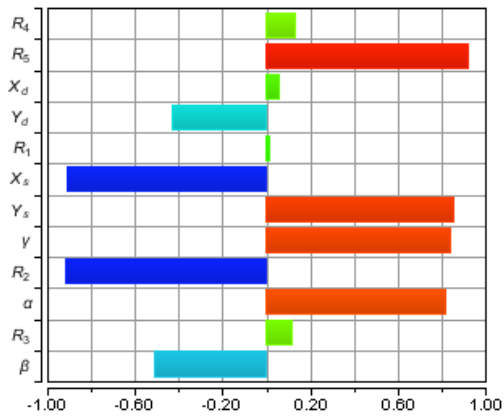
## A.2 Graphische Darstellungen zum Anwendungsbeispiel



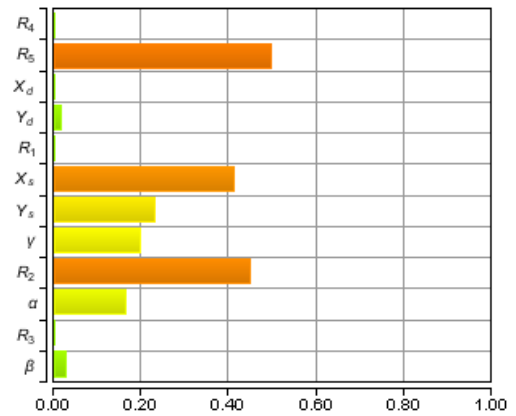
(a) Pearson'sche Korrelationskoeffizienten



(b) Kendall'sche Korrelationskoeffizienten

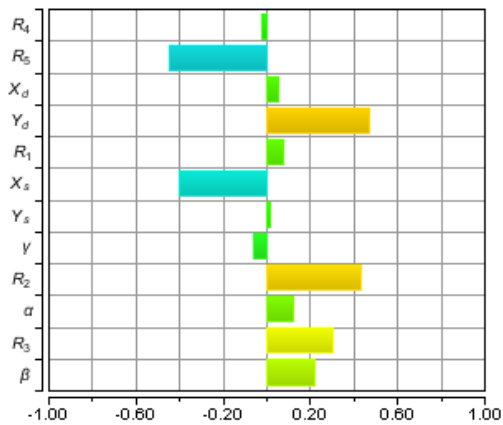


(c) Partielle Korrelationskoeffizienten

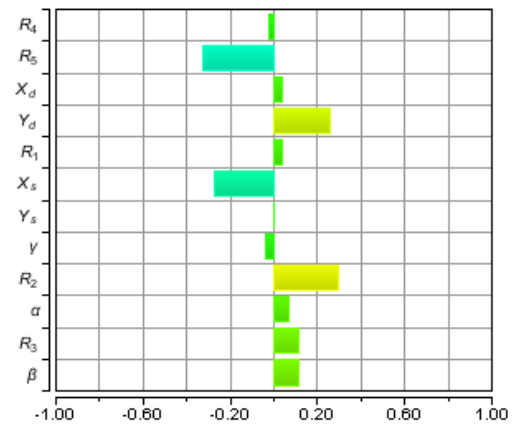


(d) Koeffizienten relativer Wichtigkeit

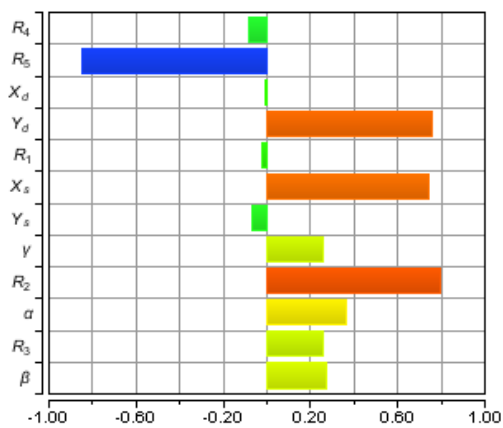
Abb. A.1: Korrelationskoeffizienten zwischen der Zielgröße V und allen Parametern



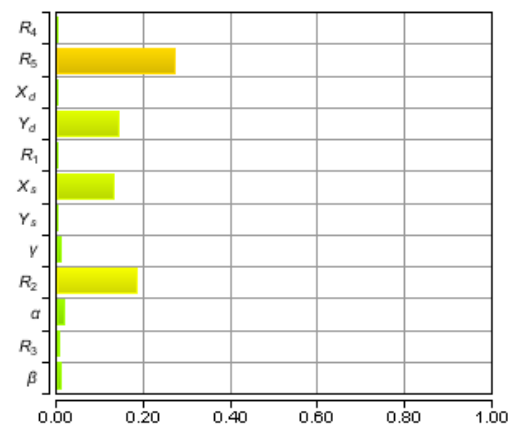
(a) Pearson'sche Korrelationskoeffizienten



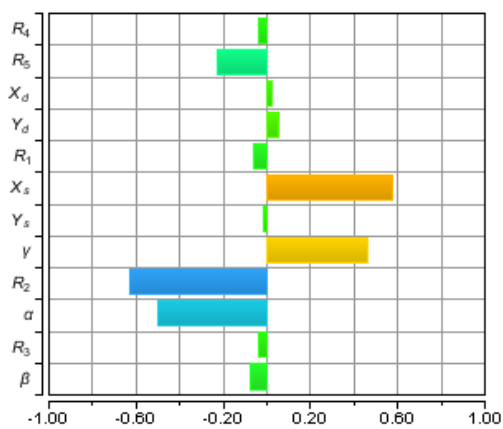
(b) Kendall'sche Korrelationskoeffizienten



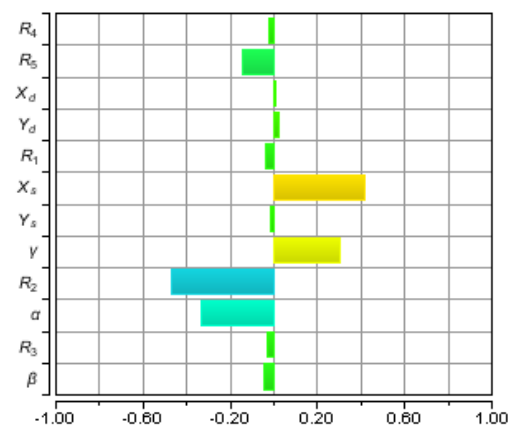
(c) Partielle Korrelationskoeffizienten



(d) Koeffizienten relativer Wichtigkeit

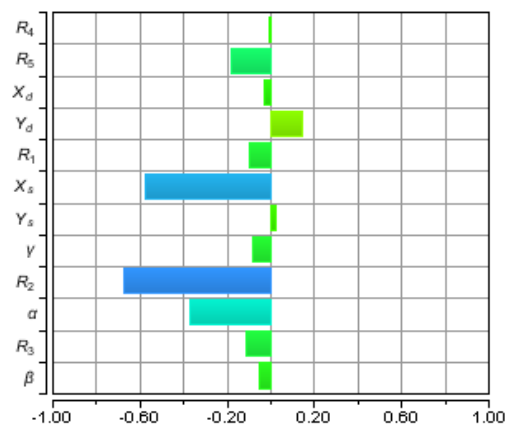
Abb. A.2: Korrelationskoeffizienten zwischen der Zielgröße  $V_v$  und allen Parametern

(a) Pearson'sche Korrelationskoeffizienten

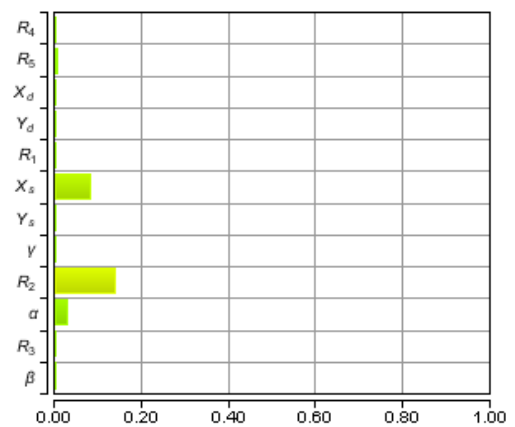


(b) Kendall'sche Korrelationskoeffizienten

Abb. A.3: Korrelationskoeffizienten zwischen der Zielgröße T und allen Parametern

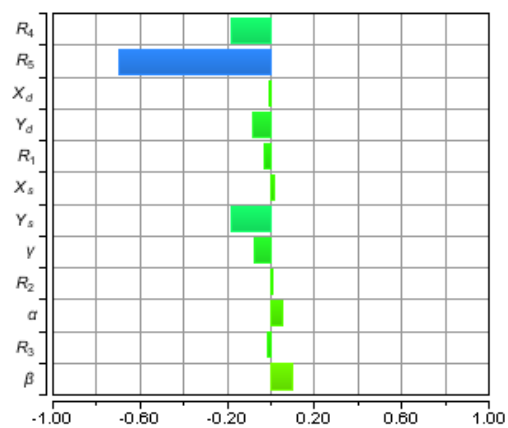


(c) Partielle Korrelationskoeffizienten

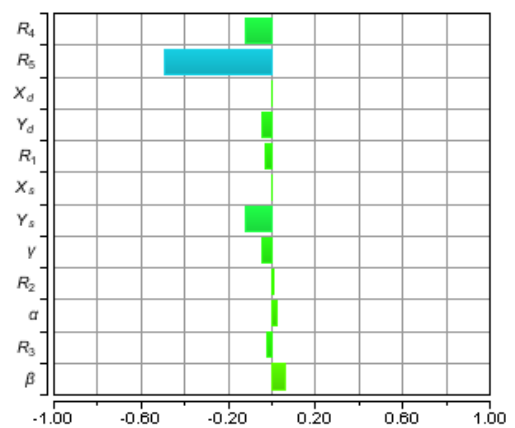


(d) Koeffizienten relativer Wichtigkeit

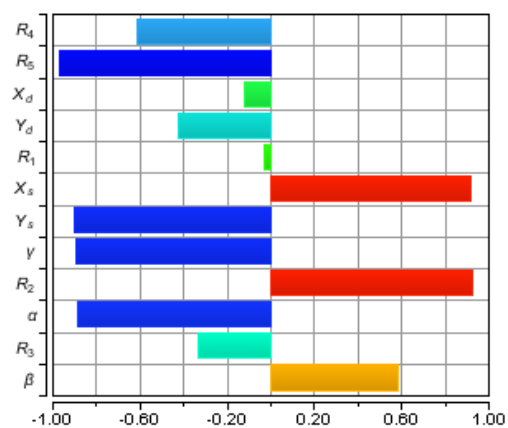
Abb. A.3: Korrelationskoeffizienten zwischen der Zielgröße T und allen Parametern



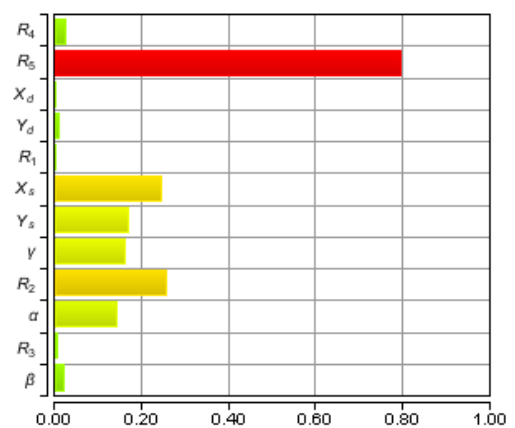
(a) Pearson'sche Korrelationskoeffizienten



(b) Kendall'sche Korrelationskoeffizienten

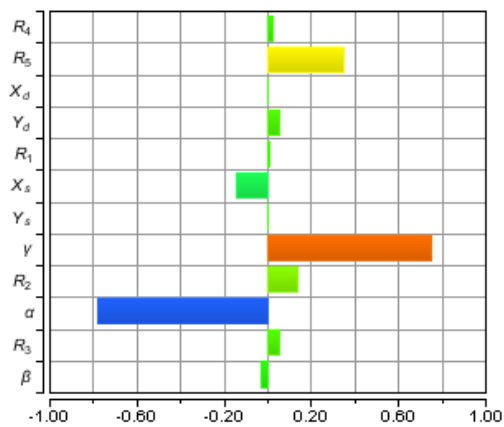


(c) Partielle Korrelationskoeffizienten

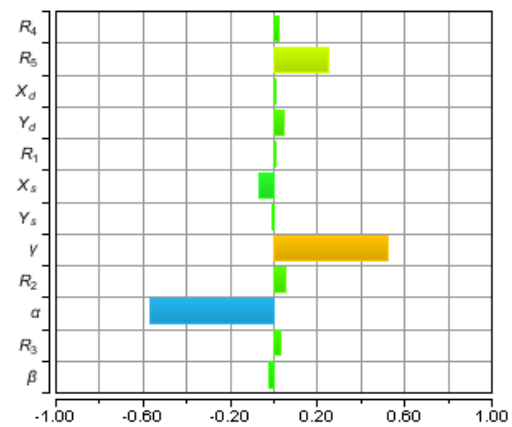


(d) Koeffizienten relativer Wichtigkeit

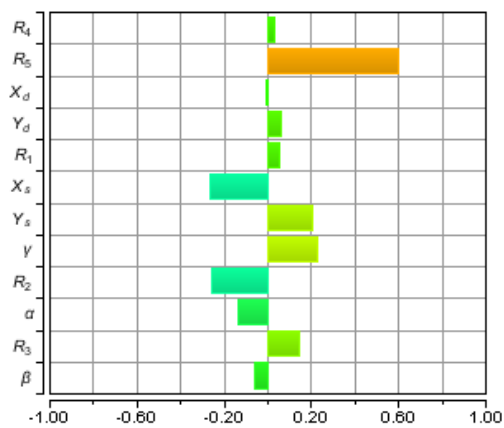
Abb. A.4: Korrelationskoeffizienten zwischen der Zielgröße  $P_r$  und allen Parametern



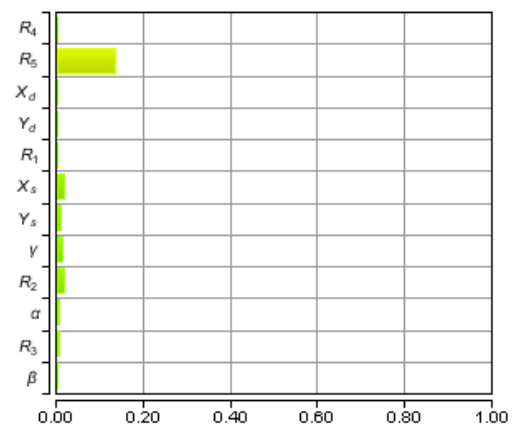
(a) Pearson'sche Korrelationskoeffizienten



(b) Kendall'sche Korrelationskoeffizienten



(c) Partielle Korrelationskoeffizienten



(d) Koeffizienten relativer Wichtigkeit

Abb. A.5: Korrelationskoeffizienten zwischen der Zielgröße U und allen Parametern

# Literaturverzeichnis

- [1] AHN, CHANG W.: *Advances in Evolutionary Algorithms*. Springer Verlag, 2006.
- [2] BENKER, HANS: *Mathematische Optimierung mit Computeralgebrasystemen*. Springer Verlag, 2003.
- [3] GÖHLER, WILHELM: *Formelsammlung Höhere Mathematik*. Verlag Harri Deutsch, 2005.
- [4] HARTUNG, JOACHIM; ELPELT, BÄRBEL: *Multivariate Statistik*. Oldenbourg Verlag, 1992.
- [5] HARTUNG, JOACHIM: *Statistik*. Oldenbourg Verlag, 2005.
- [6] HÄMMERLIN, GÜNTHER; HOFFMANN, KARL-HEINZ: *Numerische Mathematik*. Springer Verlag, 1989.
- [7] IRLE, ALBRECHT: *Wahrscheinlichkeitstheorie und Statistik*. Teubner Verlag Stuttgart, 2001.
- [8] KÖNNING, MARKUS: *Optimierung und Robustheitsbewertung in der Simulation mechanischer Systeme*. Dissertation, Bauhaus - Universität Weimar, 2007.
- [9] KOHN, WOLFGANG: *Statistik - Datenanalyse und Wahrscheinlichkeitsrechnung*. Springer Verlag, 2005.
- [10] KUX, STEFFEN; PARSCHE, UWE: *Effektive Optimierungsverfahren im Auslegungsprozess für Kettentriebe*. MTZ, 70. Jahrgang: 58 – 65, 01/2009.
- [11] MEYER, MICHAEL: *Entwicklung eines Expertensystems zur Unterstützung der Auswahl und Parametrisierung von multikriteriellen Optimierungsverfahren für Aufgabestellungen in der Motorenentwicklung*. Diplomarbeit, Technische Universität Chemnitz, 2009.
- [12] RÖBER, MARCEL: *Multikriterielle Optimierungsverfahren für rechenzeitintensive technische Aufgabestellungen*. Diplomarbeit, Technische Universität Chemnitz, 2010.
- [13] SACHS, LOTHAR: *Angewandte Statistik*. Springer Verlag, 2003.
- [14] SCHWARZ, HANS-RUDOLF: *Numerische Mathematik*. Teubner Verlag Stuttgart, 1993.
- [15] SEN, ASHISH; SRIVASTAVA, MUNI: *Regression Analysis*. Springer Verlag, 1990.

[16] <http://www.it.lut.fi/ip/evo/functions/node27.html>, verfügbar am 22.07.10.

[17] <http://www.versuchsmethoden.de/Taguchi.htm>, verfügbar am 30.07.10.

# Erklärung

Hiermit erkläre ich, dass ich die vorliegende Arbeit selbstständig und nur unter Verwendung der angegebenen Literatur und Hilfsmittel angefertigt habe.

Stellen, die wörtlich oder sinngemäß aus Quellen entnommen wurden, sind als solche kenntlich gemacht.

Diese Arbeit wurde in gleicher oder ähnlicher Form noch keiner anderen Prüfungsbehörde vorgelegt.

Chemnitz, 30. August 2010